Robust Predictions with Bounded Information*

Benjamin Brooks Songzi Du Alexander Haberman

July 15, 2025

Abstract

We study robust predictions for Bayesian games, where players have limited information about both the state of the world and others' information. These limits restrict which joint distributions between players' actions and the state are *feasible*, and they also affect the form of the players' *obedience constraints*. We have three main results. First, we characterize when a constraint on the information only restricts what is feasible, without affecting obedience. Second, we characterize exactly which feasibility constraints can arise from constraints on information. A leading example is an upper bound on the divergence between players' joint information about the state and the prior. Third, for a class of linear games, we show that divergence bounds on the players' joint information are equivalent to assuming that there is a set of possible prior distributions of the state, which are mean-preserving contractions of the true prior. We apply the theory to coordinated attack games, auctions, and informationally-robust mechanism design.

Keywords: Robust predictions, Bayes correlated equilibrium, information design, mechanism design, auctions.

JEL Classification: C72, D44, D82, D83.

---

*Brooks: Department of Economics, University of Chicago, babrooks@uchicago.edu; Du: Department of Economics, University of California—San Diego, sodu@ucsd.edu; Haberman: Department of Economics, Stanford University, ahabermn@stanford.edu. We have benefitted tremendously from feedback from Alex Bloedel, Emir Kamenica, and Jay Lu, as well as a number of seminar audiences. Tyler Patterson provided excellent research assistance. Brooks and Du gratefully acknowledge support from NSF #2215259.

1 Introduction

The standard approach to Bayesian games is to study equilibrium under a particular information structure. A growing literature on informationally-robust predictions has instead analyzed the range of possible outcomes across *all* information structures within a large class, often with restrictions only on the distribution of payoff-relevant states. This modeling approach addresses two issues: First, conclusions about behavior will necessarily be less dependent on stylized assumptions about information. Second, as long as the set of information structures is sufficiently rich, then the problem can be reformulated as the analysis of *Bayes correlated equilibria* (BCE) (Bergemann and Morris, 2016), which are joint distributions of actions and states (called *outcomes*) that satisfy a system of linear inequalities.

A weakness of the robust theory is that the set of information structures is so large. In particular, there are no limits to how well informed players might be about the state and others' information.[1] Depending on the application and the information structure, the players' information may be implausibly precise.

To put a finer point on it, take, for example, wildcatters bidding for an oil tract in a first-price auction. In the robust model, we are agnostic about the precise form of the bidders' information. Under the standard assumption that the bidders have a common value, Bergemann et al. (2017) derived the information and equilibrium that minimize expected revenue, and showed that it has the feature that by pooling their information, the bidders would know *precisely* the value of the oil tract. Do we really think that the bidders can resolve all of the uncertainty? Or would a more reasonable model be one in which there will be some residual uncertainty about values (because, for example, the bidders cannot perfectly infer the geology deep underground), but of a form that the seller cannot precisely predict ex ante?

In this paper, we address this issue by proposing and operationalizing a refinement of BCE, which we call *feasibility constrained BCE* (FCBCE). This refinement is a robust prediction for play across all equilibria and for a restricted set of information structures. FCBCE is formulated with two objectives in mind: To incorporate a flexible class of constraints on information that are economically meaningful, while also preserving the analytical tractability of BCE. Regarding the economic content, the restrictions on information that are implicit in FCBCE can be interpreted as *upper bounds* on the players' information.

[1] The formulation of BCE introduced in Bergemann and Morris (2016) incorporates lower bounds on the players' information, in what they refer to as the *individual sufficiency* order on information structures. For our main results, we assume that the lower bound is degenerate. In Section 7, we incorporate the lower bound into our theory.

Regarding tractability, we show that this class of restrictions on the information structure only constrain which outcomes are *feasible*, and do not affect the form of the obedience constraints. This allows us to effectively "solve out" the information structure, and reformulate the restriction on information as a feasibility constraint, thereby preserving the analytical tractability present in the baseline model. These ideas are explained in more detail in the following paragraphs.

To motivate FCBCE, consider first the case of a single player, for which information can be described as an experiment in the sense of Blackwell. Suppose we wish to put an upper bound on how much this player knows. A natural way to do so would be to fix an "upper bound" experiment, and let admissible experiments be those that are less informative than the upper bound, in the Blackwell order. Equivalently, the player's information is assumed to be a garbling of the upper bound, whereby each signal is replaced with a noisy observation, and the noise does not depend on the state. The single upper bounding experiment is quite special. More generally, we could fix a "frontier" of most informative experiments, and suppose that admissible experiments are those that are garblings of one of the experiments on the frontier. An equivalent way of formulating such a set of experiments is to require that the set of experiments is *complete* under garbling: If an experiment is admissible, then so is every garbling of that signal.

Our theory generalizes this notion of "garbling completeness" to multiple players, where instead of experiments, we use information structures. The correct generalization of garbling is an *individual garbling*: Each player's signal is replaced with a noisy observation, where the noise does not depend on the state or on other players' signals. We also generalize the notion of completeness. A subtlety that arises with two or more players is that it may be possible for the players to coordinate on ignoring some aspect of their signals in a manner that is self reinforcing. An extreme case is when the information structure is entirely uninformative about the state (and is actually a pure correlation device). In that case, a player has no reason to use the correlation device when others are ignoring it. More generally, we say that an individual garbling is *coordinated* if players would garble their information voluntarily, provided others do so as well. A set of information structures is *individual garbling complete* if every individual garbling of something in the set is also a coordinated individual garbling of something in the set. This is a set of information structures that are below a "frontier" of most informative information structures, analogous to what we previously described with a single player.

For any set of admissible information structures, one could consider the corresponding robust prediction, where we range over information structures in the set and equilibria. When the admissible set contains all information structures, as in the baseline model, there is a "revelation principle" for information design that any outcome which is induced by some information structure and equilibrium can also be induced by a *direct* information structure, in which each player's signal is a recommended action, and the *obedient* equilibrium, in which each player's action is equal to their signal. BCE are precisely those direct information structures for which obedient strategies are an equilibrium. However, this result crucially relies on the fact that all information structures are feasible, and it need not hold with an arbitrary restriction on information (and it is certainly not true in the degenerate case when the set contains a single information structure). However, our Theorem 1 shows that there is an analogous revelation principle when the set of admissible information structures is individual garbling complete: For every game, the induced equilibrium outcomes are those which satisfy the standard obedience constraints and are *feasible*, meaning that they can be induced by some strategies of the players for some information structure in the set. Moreover, the converse is also true: Given a set of admissible information structures, if it is the case that for every game the equilibrium outcomes are those that are feasible and obedient, then the set of information structures is individual garbling complete.

Thus, any restriction on information that is an individual garbling complete can be equivalently formulated as a restriction on which outcomes are feasible. This is an enormous simplification, since it allows us to treat the information structure as implicit and conduct all of our analysis on outcomes, using direct information structures and obedient equilibria. Importantly, however, a set of feasible outcomes that is derived from a set of information structures must take on a special form. In fact, an admissible set of information structures induces a whole *feasibility correspondence*, which maps games into sets of feasible outcomes. Understanding the whole correspondence is essential for applying the theory to mechanism design, where we optimize over games.

Theorem 2 shows that a feasibility correspondence is induced by some set of information structures if and only if the correspondence satisfies an analogous notion of individual garbling completeness: starting from an outcome that is feasible under one game, any individual garbling of that outcome to a second game's action space must be feasible for the second game. Theorem 2 therefore allows us to completely solve out information, and operationalize upper bounds on information in terms of what is feasible. This finally brings us to the definition of FCBCE: The set of outcomes that are obedient and feasible relative to a fixed *individual garbling complete feasibility correspondence.*

An additional source of tractability for the baseline model is that the set of BCE is convex. For FCBCE, this will be the case whenever the feasibility correspondence is convex valued. Theorem 3 shows that this is equivalent to another property of the set of information structures, that it is *public randomization complete*.

The remainder of this paper operationalizes FCBCE with a particular parametric class of feasibility correspondences, applied to robust predictions in games and mechanism design. This parametric class is defined by an upper bound on how much information the players collectively have about the state. Information is measured by an f-divergence, for some convex function f. Special cases include total variation distance or Kullback-Leibler divergence (i.e., mutual information). When the upper bound is zero, this forces the action profile and the state to be independent, so that players effectively have no information about the state. When the bound is sufficiently large, there is no restriction on the correlation between action profiles and states. Theorem 4 shows that for any f-divergence, the resulting feasibility correspondence is both individual garbling complete and convex valued. Moreover, we provide a structural characterization of extreme points of the set of f-divergence FCBCE, in the special case where the players' utilities are linear in the state: the extreme points coincide with extreme points of the set of (unconstrained) BCE, but where the distribution of the state is a mean-preserving contraction of the true prior.

We apply FCBCE to a coordinated attack problem, a first-price auction, and informationally robust optimal auction design. For the coordinated attack problem, we compute the BCE that maximize the probability of an attack. A qualitative message is that the tighter is the constraint on players' information, the more weight the extremal BCE place on actions where the agents attack less, since these actions are easier to incentivize when there is less available information about the state. For the first-price auction, we compute the BCE that minimizes expected revenue subject to a divergence constraint. Here, we confirm that the optimal BCE has the same form as that described by Bergemann, Brooks, and Morris (2017): revenue is minimized when the bidders receive independent signals, the interim expected value is the maximum of the signals, and the prior is a mean-preserving spread of the highest signal. The final application concerns the design of auctions that maximize minimum revenue with pure common values, as in Brooks and Du (2021), but with the f-divergence constraint. Again, we find that the optimal auction has the same form as in Brooks and Du (2021), but for a contracted prior.

The penultimate section of the paper enriches the definition of FCBCE to incorporate lower bounds on information, in the form of a base information structure as in Bergemann and Morris (2016). We informally discuss how our main results would extend, and we also describe how f-divergence bounded feasibility correspondences could be used to interpolate between BCE and other existing solution concepts for Bayesian games, including the Bayesian solution (Forges, 1993), belief-invariant BCE (Liu, 2009), as well as a novel solution concept that we call *group belief invariant BCE*. That section concludes with simulations of a first-price auction with private values, as in Bergemann, Brooks, and Morris (2017, 2021).

In addition to the aforementioned work, our analysis also relates to other studies of relations on information structures. Lehrer, Rosenberg, and Shmaya (2013) characterize when two information structures induce the same set of equilibrium outcomes, under various equilibrium concepts. One of their results is that two information structures have the same Bayes Nash equilibrium outcomes for all games if and only if they are individual garblings of one another. Gossner (2000) asks when one information structure has *more* Bayes Nash equilibrium outcomes than another information structure, for every game. He argues that this is equivalent to the coordinated individual garbling relation. (Gossner refers to a coordinated individual garbling as a "faithful reproduction.") There are important technical differences between his results and ours, and in particular, Gossner relies on infinite games in order to provide his characterization, whereas we restrict attention to finite games. These differences will be discussed in greater detail in Section 3.

The rest of this paper is organized as follows. Section 2 describes our model. Section 3 contains our main epistemic results on FCBCE. Section 4 presents our results on f-divergence FCBCE. Section 5 describes our results on linear games, and Section 6 applies those results in turn to maxmin mechanism design. Section 7 informally discusses the addition of the lower bound on information. Section 8 concludes the paper. Appendix A contains omitted proofs, and Append B describes the connection between our work and other research on binary relations on information structures.

2 Preliminary notation

There is a finite set of players indexed by $i = 1, \ldots, N$. Preferences depend on a state of the world $\theta \in \Theta$, with Θ also finite. The sets of players and states are held fixed throughout our analysis.

The players' private information is described by a (common prior) *information structure*, which consists of the following: a *signal space*, which is a finite product set $S = \prod_{i=1,\ldots,N} S_i$; and a joint distribution $\sigma \in \Delta(S \times \Theta)$. An information structure is denoted $I = (S, \sigma)$. We identify sets of signals with subsets of the integers, so that the set of all information structures is well defined.

A *prior* over the state is just a distribution $\mu \in \Delta(\Theta)$. Given a prior μ, we say that $I = (S, \sigma)$ is *consistent with* μ if the marginal of σ on Θ is μ.

An *action space* is a finite product set of the form $A = \prod_{i=1,\ldots,N} A_i$. As with signals, we identify sets of actions with subsets of the integers, so that the set of all action spaces is well defined.

The players interact through a *game structure* (also variously known as a game form or a base game), which consists of an action space A and, for each player i, an expected utility index $u_i : A \times \Theta \to \mathbb{R}$. The game structure is denoted by $G = (A, u)$.

A *Bayesian game* is a pair (I, G) of an information structure and a game structure. A (behavioral) strategy for player i is a mapping $b_i : S_i \to \Delta(A_i)$. A profile of strategies $b = (b_1, \ldots, b_N)$ is identified with the mapping $b : S \to \Delta(A)$, where $b(a|s) = \prod_{i=1,\ldots,N} b_i(a_i|s_i)$. The set of strategies of player i is denoted $B_i(S, A)$, and the of strategy profiles is $B(S, A)$. Given $b \in B(S, A)$, player i's expected utility is

$$U_i(b; I, G) = \sum_{\theta \in \Theta} \sum_{s \in S} \sum_{a \in A} u_i(a, \theta) b(a|s) \sigma(s, \theta).$$

The profile b is a *(Bayes Nash) equilibrium* if $U_i(b; I, G) \geqslant U_i(b_i', b_{-i}; I, G)$ for all i and $b_i' \in B_i(S, A)$.

Given an action space A, an *outcome* is a distribution $\phi \in \Delta(A \times \Theta)$. An information structure $I = (S, \sigma)$ and strategies $b \in B(S, A)$ *induce* an outcome ϕ defined by

$$\phi(a, \theta) = \sum_{s \in S} b(a|s) \sigma(s, \theta).$$

Given I and A, the set of *feasible outcomes* $F_I(A)$ is the set of outcomes induced by $I = (S, \sigma)$ and some $b \in B(S, A)$. Given a game (I, G), a $\phi \in \Delta(A \times \Theta)$ is an *equilibrium outcome* if there exists an equilibrium b of (I, G) such that (I, b) induce ϕ. The set of equilibrium outcomes is $E_I(G)$.

Given a set of information structures $\mathcal{I}$, we define $E_{\mathcal{I}}(G) = \cup_{I \in \mathcal{I}} E_I(G)$ and $F_{\mathcal{I}}(A) = \cup_{I \in \mathcal{I}} F_I(A)$.

Fix a game structure $G = (A, u)$. Following Bergemann and Morris (2013, 2016) and Bergemann, Brooks, and Morris (2022), we say that the outcome $\phi \in \Delta(A \times \Theta)$ is a *Bayes correlated equilibrium (BCE)* of G if for all i, a_i, and a_i', the following inequality holds:

$$\sum_{a_{-i} \in A_{-i}, \theta \in \Theta} \phi(a_i, a_{-i}, \theta)(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)) \geqslant 0. \tag{1}$$

The inequalities (1) are referred to as *obedience constraints*, and we also call a BCE outcome an *obedient outcome*. We write BCE(G) for the set of BCE of G. It follows immediately from Theorem 1 of Bergemann and Morris (2016) that $\phi \in E_I(G)$ for some I if and only if $\phi \in \text{BCE}(G)$.[2]

3 Individual garbling completeness and feasibility constraints

In this section, we provide our main epistemic characterizations of feasibility-constrained BCE: We formulate the notion of individual garbling completeness of a set of information structures, and we show that this condition is necessary and sufficient for the implied restriction on equilibrium outcomes to only operate through feasibility. We also characterize precisely those feasibility correspondences which can be induced by individual garbling complete sets of information structures. Finally, we give further conditions that characterize when the set of FCBCE is also convex, namely, that the set of information structures is public randomization complete.

3.1 Individual garbling completeness

Given information structures $I = (S, \sigma)$ and $I' = (S', \sigma')$, we say that I is an *individual garbling* of I' if there exist mappings $b_i : S'_i \to \Delta(S_i)$ for each i such that

$$\sigma(s, \theta) = \sum_{s' \in S'} b(s|s')\sigma'(s', \theta)$$

for all (s, θ). In a slight abuse of terminology, we also refer to b as the individual garbling (from I' to I). We note for future reference that the individual garbling relation is transitive, so that if I is an individual garbling of I' and I' is an individual garbling of I'', then I is an individual garbling of I''.

We further say that I is a *coordinated individual garbling* of I' if I is an individual garbling of I' via a mapping b, and moreover, for every s_i and s'_i such that $\sigma'(\{s'_i\} \times S'_{-i} \times \Theta) > 0$ and $b_i(s_i|s'_i) > 0$, we have

$$\sigma(s_{-i}, \theta \mid s_i) = \sum_{s'_{-i} \in S'_{-i}} \sigma'(s'_{-i}, \theta \mid s'_i) \prod_{j \neq i} b_j(s_j|s'_j) \tag{2}$$

[2]The definition of BCE given in Bergemann and Morris (2013, 2016) also imposes a fixed marginal distribution over θ. However, their result immediately extends to the case where the marginal on θ is allowed to "float," as we have done here. This is the version of BCE used in Bergemann, Brooks, and Morris (2022).

for every s_{-i} and θ, where $\sigma(\cdot,\cdot|s_i)$ denotes the conditional belief given s_i, updating from the prior σ. In words, if the signal s'_i is garbled to s_i with positive probability, then conditional on s_i, beliefs about the state and others' garbled signals (θ, s_{-i}) do not depend on s'_i. Gossner (2000) refers to a coordinated individual garbling as a "faithful reproduction," in that if I is a coordinated individual garbling of I', then starting from I', it is possible for the players to "reproduce" I in equilibrium, by independently garbling their own information.

We will illustrate these definitions with a simple example. Suppose that $N = 2$ and $\Theta = \{0,1\}$. Consider the information structure $I = (S, \sigma)$ where $S_i = \{0,1\}$ and the joint distribution of signals and states σ is

$\theta = 0$			$\theta = 1$		
$s_1 \backslash s_2$	0	1	$s_1 \backslash s_2$	0	1
0	1/4	0	0	0	1/4
1	0	1/4	1	1/4	0

In other words, all signal profiles have probability 1/4, and the state is equal to the parity of the sum of the signals.

Next, consider the "no information" structure $I' = (S'_i, \sigma')$ where $S'_i = \{\varnothing\}$ for $i = 1,2$, and $\sigma'(\varnothing, \varnothing, \theta) = 1/2$ for each $\theta \in \Theta$. No information is clearly an individual garbling of I, where $b_i(\varnothing|s_i) = 1$ for all i. It is also a coordinated individual garbling. Conditional on $s'_i = \varnothing$, both $(s'_j, \theta) = (\varnothing, 0)$ and $(\varnothing, 1)$ are equally likely. Moreover, these are the same beliefs about (s'_j, θ) that player i would have conditional on any $s_i \in S_i$. In effect, the players' signals are individually uninformative about the state. So if both players ignore their signals, then neither will have an incentive to use their signals.

For an example of an information structure $I'' = (S'', \sigma'')$ that is individual garbling of I and *not* a coordinated individual garbling, take $S'' = S$, and σ'' is given by the following table:

$\theta = 0$			$\theta = 1$		
$s''_1 \backslash s''_2$	0	1	$s''_1 \backslash s''_2$	0	1
0	3/16	0	0	0	3/16
1	1/16	1/4	1	1/4	1/16

This information structure can be obtained from I with the individual garbling $b_1(0|0) = 3/4$, $b_1(1|0) = 1/4$, $b_1(1|1) = 1$, $b_1(0|1) = 0$, and $b_2(s_2|s_2) = 1$ for all s_2. But it is not a coordinated individual garbling: Conditional on $s''_1 = 1$, all (s''_2, θ) have positive probability. But conditional on $s_1 = 0$ (which garbles to $s''_i = 1$), there is zero probability that $(s''_2, \theta) = (0,1)$ and $(1,0)$. So, if only player 1 adds noise to their signal, but player 2 continues to use theirs, then player 1 would have an incentive to look at their ungarbled signal.

A set of information structures $\mathcal{I}$ is *individual garbling complete* if every information structure that is an individual garbling of an element of $\mathcal{I}$ is also a coordinated individual garbling of an element of $\mathcal{I}$, i.e., if $I \in \mathcal{I}$ and I' is an individual garbling of I, then there exists $I'' \in \mathcal{I}$ such that I' is a coordinated individual garbling of I''.

Before proceeding, we remark on an alternative notion that naturally arises in this context. A set of information structures is *closed under individual garbling* if for any $I \in \mathcal{I}$ and I' that is an individual garbling of I, $I' \in \mathcal{I}$. This condition is stronger than individual garbling completeness, in that the latter would only require that I' is a coordinated individual garbling of some information structure $I'' \in \mathcal{I}$ (and any information structure is trivially a coordinated individual garbling of itself). Thus, closed under individual garbling implies individual garbling completeness. We have focused on the weaker condition because that is the one that is useful for characterizing equilibrium predictions. Indeed, it may be that I' is just a "relabeling" of the signals in I, in which case whether I' is in $\mathcal{I}$ is irrelevant to the equilibrium outcomes that are consistent with $\mathcal{I}$ in any game.

With these definitions in hand, we can now state our first result:

Theorem 1. *$\mathcal{I}$ is individual garbling complete if and only if for every $G = (A, u)$,*

$$E_{\mathcal{I}}(G) = F_{\mathcal{I}}(A) \cap \mathrm{BCE}(G). \tag{3}$$

We can illustrate the theorem using the aforementioned information structures. Consider the game $G = (A, u)$ where $A_i = \{0, 1\}$, and

$$u_i(a, \theta) = \begin{cases} 1 & \text{if } a_1 = a_2 \text{ and } \theta = 0; \\ 1 & \text{if } a_1 \neq a_2 \text{ and } \theta = 1; \\ 0 & \text{otherwise.} \end{cases}$$

So, players want to match their actions in state 0 and mismatch in state 1.

Let us initially suppose that the set $\mathcal{I}$ consists of those information structures that are individual garblings of the information structure $I = (S, \sigma)$ constructed above. It is immediate that this set is individual garbling complete, given that every information structure is a coordinated individual garbling of itself. Moreover, any feasible and obedient outcome given $\mathcal{I}$ is also an equilibrium outcome: Suppose that ϕ is induced by some information structure $(S', \sigma') \in \mathcal{I}$ and strategies b'. Then the "direct recommendation" information structure (A, ϕ) is, by definition, an individual garbling of (S', σ'). And by hypothesis, (S', σ') is an individual garbling of (S, σ). Hence, (A, ϕ) is itself in $\mathcal{I}$. Finally, since ϕ is obedient, (A, ϕ) and the obedient strategies induce ϕ as an equilibrium outcome.

For a second example, suppose that we take $\mathcal{I}$ and produce a new set $\mathcal{I}'$ by removing all of those information structures that are "equivalent" to no information I', in the sense that the players' signals are independent of one another and of the state. (A general and precise notion of equivalence is discussed in Appendix B.) This set is still individual garbling complete: Indeed, all we have removed are the no-information structures, but as we argued above, no information is a coordinated individual garbling of I. Moreover, the set of equilibrium outcomes induced by $\mathcal{I}'$ is the same as that for $\mathcal{I}$: The only equilibrium outcomes that can be induced by no information are Nash equilibria of the ex ante game, for example playing $a = (0, 0)$ with probability one. But this is also an equilibrium under I, where both players choose $a_i = 0$ regardless of s_i.

As a final example, suppose that $\mathcal{I}''$ consists of just I. This set is not individual garbling complete, simply because there are individual garblings of I that are not coordinated individual garblings of I, such as the particular I'' constructed above. Moreover, there are feasible and obedient outcomes of G which are not equilibrium outcomes. In particular, consider the outcome induced by the "obedient" strategies on I'', in which both players play actions equal to their signals. The resulting outcome is precisely σ''. Clearly outcome σ'' is feasible under I, and one can check that it is also obedient on G. The only way to generate this outcome using I is that $b_1(1|1) = 1$ and $b_2(s_2|s_2) = 1$ for all s_2; otherwise, there would be positive probability of either $(a_1, a_2, \theta) = (0, 1, 0)$ or $(0, 0, 1)$. Hence, it must be that $b_1(0|0) = 3/4$ and $b_1(1|0) = 1/4$. But then player 1 would be strictly better off by deviating to the obedient strategy $b'_1(s_1|s_1) = 1$ for all s_1. The bottom line is that because $\mathcal{I}''$ is not individual garbling complete, we can find games for which there are feasible and obedient outcomes which are not equilibrium outcomes.

3.2 Proof of Theorem 1

We now present the proof of Theorem 1. The more technical parts of the proof will be sketched, with details in the appendix. This section can be skipped by readers who prefer to move ahead to the definition of FCBCE and our applications.

3.2.1 If

Suppose that $\mathcal{I}$ satisfies (3) for all G. Moreover, suppose that $I = (S, \sigma)$ is an individual garbling of some information structure in $\mathcal{I}$. We will prove that I is also a coordinated individual garbling of some element of $\mathcal{I}$. The proof relies on the following lemma, which is of some independent interest (as we discuss further in Appendix B).

Lemma 1. *For every $I = (S, \sigma)$, there exists a game G and an equilibrium outcome $\phi \in E_I(G)$, such that if $\phi \in E_{I'}(G)$, then I is a coordinated individual garbling of I'.*

We refer to the G in the statement of the lemma as the *separation game* for I. To see why the lemma implies the if direction of Theorem 1, suppose that I is an individual garbling of some element of $\mathcal{I}$. Let $G = (A, u)$ be the separation game for I and ϕ the equilibrium outcome, as in Lemma 1. Then clearly $\phi \in \text{BCE}(G)$, and because I is an individual garbling of something in $\mathcal{I}$, $\phi \in F_{\mathcal{I}}(A)$ as well. But because (3) holds for all games, we know that $\phi \in E_{\mathcal{I}}(G)$ as well, so $\phi \in E_{I'}(G)$ for some $I' \in \mathcal{I}$. By Lemma 1, I is a coordinated individual garbling of I'. Since I was arbitrary, we conclude that $\mathcal{I}$ is individual garbling complete.

The formal proof of Lemma 1 is in the Appendix. We will now sketch the argument. Fix an information structure $I = (S, \sigma)$. In the separation game, each player will report either a signal $s_i \in S_i$ or a "spoiler" action, which consists of a signal s_i and a direction $d \in \mathbb{R}^{S_{-i} \times \Theta}$, i.e., a direction in the space of possible beliefs about (s_{-i}, θ). The equilibrium outcome ϕ referred to in the lemma will simply be the outcome induced by the obedient strategies, i.e., each player playing an action equal to their realized signal.

The payoffs are constructed so that the obedient strategies are an equilibrium, but also so that reporting s_i is a best response to a conjecture over $\Delta(S_{-i} \times \Theta)$ if and only if the belief is precisely that of type s_i in the information structure I, which we denote by

$$\sigma(s_{-i}, \theta | s_i) = \frac{\sigma(s_i, s_{-i}, \theta)}{\sum_{s'_{-i} \in S_{-i}, \theta' \in \Theta} \sigma(s_i, s'_{-i}, \theta')}.$$

To see why these properties suffice to prove the lemma, note that if ϕ is an equilibrium outcome for $I' = (S', \sigma')$, then there are strategies b that induce ϕ as an outcome. Clearly, these strategies show that I is an individual garbling of I', and in fact, they also satisfy the belief-sufficiency property (2): If $b_i(s_i | s'_i) > 0$, but the belief at s'_i about (s_{-i}, θ) is not $\sigma(s_{-i}, \theta | s_i)$, then one of the spoiler actions is a strictly better response than s_i, which would contradict the hypothesis that b is an equilibrium. Hence, I is a coordinated individual garbling of I'.

But how are these payoffs constructed? Clearly, for the aforementioned properties to hold, it is irrelevant what the payoffs are at action profiles where more than one player takes an action that is not a reported signal. For the remaining action profiles, we construct the payoffs in two stages, first specifying payoffs if the action profile is in S, and then constructing payoffs for the spoiler actions that are strictly better responses than s_i at beliefs other than $\sigma(\cdot \mid s_i)$.

To construct payoffs for action profiles in S, we denumerate the possible interim beliefs $\sigma(\cdot \mid s_i)$ of player i under I as $\psi_i^1, \ldots, \psi_i^K \in \Delta(S_{-i} \times \Theta)$. These beliefs are arranged so that each belief is *not* in the convex hull of the ones that precede it. This means that we can find a hyperplane $\nu^k \in \mathbb{R}^{S_{-i}\times\Theta}$ that separates ψ_i^k from its predecessors in the list, by which we mean that $\nu^k \cdot (\psi_i^k - \psi_i^l) > 0$ for $l < k$. All actions s_i that correspond to the same belief ψ_i^k will be assigned the same utility u_i^k. We set u_i^1 to an arbitrary constant, and inductively set u_i^k to be

$$\begin{aligned} u_i^k(s_{-i}, \theta) = 1 + \max_{l<k} & \sum_{s'_{-i}\in S_{-i}, \theta' \in \Theta} u_i^l(s'_{-i}, \theta')\psi_i^k(s'_{-i}, \theta') \\ & - \alpha \left(\sum_{s'_{-i}\in S_{-i}, \theta' \in \Theta} \nu^k(s'_{-i}, \theta')\psi_i^k(s'_{-i}, \theta') - \nu^k(s_{-i}, \theta) \right), \end{aligned}$$

where α is a large, positive number. Note that the term involving α drops out of the expectation of u_i^k under ψ_i^k, but at any of the beliefs ψ_i^l for $l < k$, this term is large and negative. We can choose α large enough so that under a belief ψ_i^l, deviating from an action with payoffs u_i^l to an action with payoffs u_i^k is strictly suboptimal, for $l < k$. Finally, the first two terms in u_i^k (i.e., $1 + \max_{l<k} \ldots$) ensures that under a belief ψ_i^k deviating from u_i^k to u_i^l is strictly suboptimal, for $l < k$.

This construction is depicted in Figure 1, where we have (with artistic license) represented the belief space $\Delta(S_{-i} \times \Theta)$ as the x-axis, and utility is on the y axis. As k increases, the corresponding belief moves farther "out." The blue curves represent the utility hyperplanes u_i^k. Notice that at each belief ψ_i^k, its own blue line lies strictly above all of those corresponding to other beliefs.

The last step of the construction is to add the aforementioned spoiler actions, which have labels of the form (k, d), where $k = 1, \ldots, K$, and d is a direction. These directions are drawn from a set D, which has the property that linear combinations of vectors in D with non-negative weights span the whole Euclidean space that contains $\Delta(S_{-i} \times \Theta)$. For example, we can take D to be a set of basis vectors and their negatives. The utility index from (k, d) is equal to $u_i^k + \epsilon d$, where ϵ is a scalar that is sufficiently small. These utility planes are depicted as the red lines in Figure 1. This bonus is small enough so that these spoiler actions are still (weakly) suboptimal at any belief ψ_i^l, but at other beliefs, it is strictly better to take one of the spoiler actions (k, d) than it is to take an action with

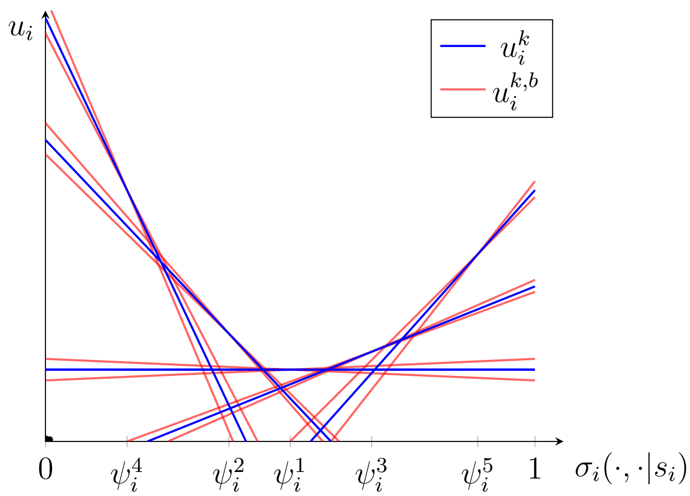

Figure 1: Constructing the separation game for I.

payoffs u_i^k. This completes the sketch of the proof of Lemma 1, and hence the proof of the if direction of Theorem 1.

3.2.2 Only if

Now suppose that $\mathcal{I}$ is individual garbling complete. Let $G = (A, u)$ and $\phi \in F_{\mathcal{I}}(A) \cap \mathrm{BCE}(G)$. Because ϕ is feasible, there is an information structure $I = (S, \sigma) \in \mathcal{I}$ and strategies in $B(S, A)$ that induce ϕ. Hence, the information structure (A, ϕ) is an individual garbling of I, and is therefore also a coordinated individual garbling of some $I' = (S', \sigma') \in \mathcal{I}$. Let $b \in B(S', A)$ be the individual garbling from (S', σ') to (A, ϕ) that satisfies (2) (replacing σ with ϕ and s_i with a_i in (2)). Clearly b induces ϕ. We claim that b is an equilibrium of (I', G). To see this, note that for every a_i and $s'_i \in S'_i$ such that $b_i(a_i|s'_i) > 0$ and s'_i has a positive probability under σ', (2) is satisfied, and therefore

$$
\begin{aligned}
&\sum_{a_{-i}\in A_{-i}, s'_{-i}\in S'_{-i}, \theta\in\Theta} \prod_{j\neq i} b_j(a_j|s'_j)\sigma'(s'_{-i}, \theta \mid s'_i) u_i(a'_i, a_{-i}, \theta) \\
&= \sum_{a_{-i}\in A_{-i}, \theta\in\Theta} \phi(a_{-i}, \theta \mid a_i) u_i(a'_i, a_{-i}, \theta) \\
&\leqslant \sum_{a_{-i}\in A_{-i}, \theta\in\Theta} \phi(a_{-i}, \theta \mid a_i) u_i(a_i, a_{-i}, \theta) \\
&= \sum_{a_{-i}\in A_{-i}, s'_{-i}\in S'_{-i}, \theta\in\Theta} \prod_{j\neq i} b_j(a_j|s'_j)\sigma'(s'_{-i}, \theta \mid s'_i) u_i(a_i, a_{-i}, \theta),
\end{aligned}
$$

where the inequality follows the fact that $\phi \in \mathrm{BCE}(G)$, and $\phi(a_{-i}, \theta | a_i)$ denotes the conditional distribution given a_i and the prior ϕ. We conclude that $\phi \in E_{I'}(G) \subseteq E_{\mathcal{I}}(G)$, as desired.

3.3 Feasibility correspondences

We now characterize the class of feasibility constraints that can be derived from some restricted set of information structures that is individual garbling complete. A *feasibility correspondence* is a function that maps each product set of action profiles A into a subset of $\Delta(A \times \Theta)$. We extend the notion of individual garblings to outcomes by associating each outcome $\phi \in \Delta(A \times \Theta)$ with its direct recommendation information structure (A, ϕ).

The feasibility correspondence F is *individual garbling complete* if for every A, A', and $\phi \in F(A)$, if (A', ϕ') is an individual garbling of (A, ϕ), then $\phi' \in F(A')$. Our next result is:

Theorem 2. *F is individual garbling complete if and only if $F = F_{\mathcal{I}}$ for some $\mathcal{I}$ that is individual garbling complete.*

This result will follow from three lemmas.

Lemma 2. *For any $\mathcal{I}$, $F_{\mathcal{I}}$ is individual garbling complete.*

Proof. Let $\phi \in F_{\mathcal{I}}(A)$. Then there exists an $I = (S, \sigma)$ and $b \in B(S, A)$ that induce ϕ. Now suppose that $\phi' \in \Delta(A' \times \Theta)$ is an individual garbling of ϕ, with the garbling itself being $b' \in B(A, A')$. Consider the strategies $\widehat{b} \in B(S, A')$ defined by

$$\widehat{b}_i(a'_i | s_i) = \sum_{a_i \in A_i} b'_i(a'_i | a_i) b_i(a_i | s_i).$$

Then the outcome induced by I and $\widehat{b}$ is

$$\begin{aligned}
\widehat{\phi}(a', \theta) &= \sum_{s \in S} \widehat{b}(a' | s) \sigma(s, \theta) \\
&= \sum_{s \in S, a \in A} b'(a' | a) b(a | s) \sigma(s, \theta) \\
&= \sum_{a \in A} b'(a' | a) \phi(a, \theta) \\
&= \phi'(a', \theta),
\end{aligned}$$

so $\phi' \in F_{\mathcal{I}}(A')$, as desired. $\square$

Given a feasibility correspondence F, let $\mathcal{I}_F$ be the corresponding set of "direct recommendation" information structures of the form (A, ϕ) for $\phi \in F(A)$.

Lemma 3. *If F is individual garbling complete, then $F = F_{\mathcal{I}_F}$.*

Proof. For any $\phi \in F(A)$, $(A, \phi) \in \mathcal{I}_F$. Moreover, (A, ϕ) together with the obedient strategies in $B(A, A)$ induce ϕ, and hence $\phi \in F_{\mathcal{I}_F}(A)$. This proves that $F \subseteq F_{\mathcal{I}_F}$ (and this is always true regardless of whether F is individual garbling complete).

Conversely, if $\phi \in F_{\mathcal{I}_F}(A)$, then there is an $(A', \phi') \in \mathcal{I}_F$ and strategies $b \in B(A', A)$ such that (A', ϕ') and b induce ϕ. Thus, the information structure (A, ϕ) is an individual garbling of (A', ϕ'), and according to our definition ϕ is an individual garbling of ϕ'. From the definition of $\mathcal{I}_F$, we know that $\phi' \in F(A')$, and hence $\phi \in F(A)$ by individual garbling completeness. This proves that $F_{\mathcal{I}_F} \subseteq F$, and we are done. $\square$

Lemma 4. *If F is individual garbling complete, then $\mathcal{I}_F$ is individual garbling complete.*

Proof. Suppose that F is individual garbling complete. Let $(A, \phi) \in \mathcal{I}_F$ and let (S, σ) be an individual garbling of (A, ϕ). Consider the action space $A' = S$ and the outcome $\phi' = \sigma$. Then clearly, the outcome ϕ' is an individual garbling of ϕ, and so by individual garbling completeness, we have $\phi' \in F(A')$, so that $(A', \phi') \in \mathcal{I}_F$. Thus, $\mathcal{I}_F$ is individual garbling complete. $\square$

Proof of Theorem 2. By Lemma 2, if F is induced by some $\mathcal{I}$, then F is individual garbling complete, whether or not $\mathcal{I}$ is itself individual garbling complete.

If F is individual garbling complete, then by Lemma 3, it is induced by $\mathcal{I}_F$, and by Lemma 4, $\mathcal{I}_F$ is individual garbling complete, so F is induced by a set of information structures that is individual garbling complete. $\square$

We note that the admissible set $\mathcal{I}_F$ is not only individual garbling complete; it is also closed under individual garblings. We therefore obtain the following proposition:

Proposition 1. *F is individual garbling complete if and only if $F = F_{\mathcal{I}}$ for some $\mathcal{I}$ that is closed under individual garbling.*

3.4 Feasibility constrained BCE

This brings us to the definition of our refinement. Given an individual garbling complete feasibility correspondence F, the associated set of *feasibility constrained Bayes correlated equilibria* (FCBCE) of a game structure $G = (A, u)$ is the set

$$BCE_F(G) \equiv F(A) \cap BCE(G).$$

By Theorems 1 and 2, we have the following:

Corollary 1. *$\mathcal{I}$ is individual garbling complete if and only if there is an individual garbling complete feasibility correspondence F such that for all G, $BCE_F(G)$ is the set of equilibrium outcomes $E_{\mathcal{I}}(G)$.*

3.5 Public randomization completeness and convexity

As a final task for this section, we address the question: under what conditions on $\mathcal{I}$ is the induced feasibility correspondence convex valued? Given information structures $I = (S, \sigma)$ and $I' = (S', \sigma')$, and $\alpha \in [0,1]$, we define $\alpha I + (1-\alpha)I'$ to be the information structure (S'', σ''), where $S_i'' = S_i \sqcup S_i'$, and

$$\sigma''(s,\theta) = \begin{cases} \alpha\sigma(s,\theta) & \text{if } s \in S; \\ (1-\alpha)\sigma'(s,\theta) & \text{if } s \in S'; \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $\alpha I + (1-\alpha)I'$ is an information structure in which there is *public randomization* between I and I', with weights α and $1-\alpha$ respectively. A set of information structures $\mathcal{I}$ is *public randomization complete* (PRC) if for every $I, I' \in \mathcal{I}$ and $\alpha \in [0,1]$, $\alpha I + (1-\alpha)I'$ is a coordinated individual garbling of some element of $\mathcal{I}$.

Theorem 3. *F is individual garbling complete and convex-valued if and only if $F = F_{\mathcal{I}}$ for some $\mathcal{I}$ that is individual garbling complete and public randomization complete.*

We therefore obtain the following analogue of Corollary 1:

Corollary 2. *$\mathcal{I}$ is individual garbling complete and public randomization complete if and only if there is an individual garbling complete and convex valued feasibility correspondence F such that for all G, $BCE_F(G)$ is the set of equilibrium outcomes $E_{\mathcal{I}}(G)$.*

4 f-Divergence Constrained Outcomes

In this section, we introduce a parametric class of convex-valued feasibility correspondences where the upper bound is given in terms of the players' f-information, an extension of mutual information for general f-divergences.

4.1 f-information

Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function with $f(1) = 0$.[3] If $\eta, \zeta \in \Delta(X)$ are probability distributions over a finite space X, such that η is absolutely continuous with respect to ζ, then the *f-divergence* is defined as

$$D_f(\eta \parallel \zeta) = \sum_{x \in X} \zeta(x) f\left(\frac{\eta(x)}{\zeta(x)}\right). \tag{4}$$

When $f(x) = x \log(x)$, (4) is the famous *Kullback–Leibler* (KL) divergence between η and ζ (alternatively, from η to ζ). The f-divergences therefore generalize KL-divergence to a family of dissimilarity measures between distributions. A fundamental property of f-divergences is the *data processing inequality* which has the interpretation that processing data cannot increase information (alternatively, cannot make it easier to distinguish distributions).

Proposition 2 (Polyanski and Wu (2023), Theorem 7.4)**.** *For any $\eta, \zeta \in \Delta(X)$ and transition kernel $K : X \to \Delta(X')$, let $\eta', \zeta' \in \Delta(X')$ be defined by $\eta'(x') = \sum_{x \in X} K(x'|x)\eta(x)$ and $\zeta'(x') = \sum_{x \in X} K(x'|x)\zeta(x)$, each $\forall x' \in X'$. Then for any f-divergence, D_f:*

$$D_f(\eta' \parallel \zeta') \leqslant D_f(\eta \parallel \zeta) \tag{DPI}$$

One notable application of KL-divergence is *mutual information*, which uses KL-divergence to measure the amount of information one random variable encodes in another. By replacing KL-divergence with f-divergence in the definition of mutual information, we can similarly generalize mutual information to be defined for any f-divergence. Formally, for any f-divergence, D_f, and joint distribution $\eta \in \Delta(X \times Y)$, the *$f$-information* of η is defined as $D_f(\eta \parallel \eta_X \otimes \eta_Y)$, where η_X denotes the marginal distribution over X induced by η, and $\eta_X \otimes \eta_Y$ denotes the joint distribution over $(X \times Y)$ that satisfies $\eta_X \otimes \eta_Y(x, y) = \eta_X(x)\eta_Y(y)$ for all $(x, y) \in (X \times Y)$. We continue to use this notation throughout the text.

Since f-information is just the f-divergence between a particular pair of joint distributions, the DPI also applies to f-information (see Polyanski and Wu (2023), Theorem 7.16).[4] When applied to outcomes of a game, f-information provides a flexible measure of the amount of information about the state encoded in an action profile. Such measures are a natural candidate for constraining which outcomes are feasible.

[3] Here we define $f(0) = f(0+)$, $0f\left(\frac{0}{0}\right) = 0$, and $0f\left(\frac{a}{0}\right) = \lim_{x \downarrow 0} xf\left(\frac{a}{x}\right) = af'(\infty)$ for $a > 0$.

[4] In particular for any $\eta \in \Delta(X \times Y)$ and $\widehat{K} : X \to \Delta(X')$ we set $\zeta = \eta_X \otimes \eta_Y$ and define K such that $K(x', y'|x, y) = \widehat{K}(x'|x)\mathbb{I}_{y'=y}$ to apply (DPI).

4.2 Bounding the players' joint information

For any f-divergence D_f, prior $\mu \in \Delta(\Theta)$ and $\epsilon \in \mathbb{R}_+$, we define the following feasibility correspondence:

$$F_{f,\epsilon,\mu}(A) = \{\phi \in \Delta(A \times \Theta) | \phi_\Theta = \mu, D_f(\phi \parallel \phi_A \otimes \phi_\Theta) \leqslant \epsilon\},$$

where ϕ_A is the marginal of ϕ on A and ϕ_Θ is the marginal of ϕ on Θ.

Theorem 4. *For any f-divergence D_f, prior $\mu \in \Delta(\Theta)$ and $\epsilon \in \mathbb{R}_+$, the correspondence $F_{f,\epsilon,\mu}$ is individual garbling complete and convex valued.*

Proof. Consider any (A, A'), and $\phi \in F_{f,\epsilon,\mu}(A)$. If (A', ϕ') is an individual garbling of (A, ϕ) then there exists $b : A \to \Delta(A')$ such that $\phi'(a', \theta) = \sum_{a \in A} b(a'|a)\phi(a, \theta)$. Define $\widehat{b} : (A \times \Theta) \to \Delta(A' \times \Theta)$ as the transition kernel such that $\widehat{b}(a', \theta'|a, \theta) = b(a'|a)$ if $\theta' = \theta$ and 0 otherwise. Plugging $\eta = \phi$, $\zeta = \phi_A \otimes \phi_\Theta$ and $K = \widehat{b}$ into Proposition 2, we obtain $D_f(\phi' \parallel \phi'_{A'} \otimes \phi'_\Theta) \leqslant D_f(\phi \parallel \phi_A \otimes \phi_\Theta)$. The garbling formula for ϕ' also makes it clear that $\phi'_\Theta = \phi_\Theta$ and thus $\phi' \in F_{f,\epsilon,\mu}(A')$. This proves that $F_{f,\epsilon,\mu}$ is individual garbling complete.

To prove convexity, consider $\phi^0, \phi^1 \in F_{f,\epsilon,\mu}(A)$ and $\lambda \in [0, 1]$. We will show that the mixture distribution parameterized by λ has f-information weakly less than ϵ. Combined with the fact that the mixture distribution has the same marginal μ over Θ, this proves its inclusion in $F_{f,\epsilon,\mu}(A)$ and hence convexity.

Let $\eta^\lambda, \zeta^\lambda \in \Delta(A \times \Theta \times Z)$ be defined as

$$\eta^\lambda(a, \theta, z) = \begin{cases} (1-\lambda)\phi^0(a, \theta) & \text{if } z = 0; \\ \lambda\phi^1(a, \theta) & \text{if } z = 1, \end{cases}$$

and

$$\zeta^\lambda(a, \theta, z) = \begin{cases} (1-\lambda)\phi^0_A(a)\phi^0_\Theta(\theta) & \text{if } z = 0; \\ \lambda\phi^1_A(a)\phi^1_\Theta(\theta) & \text{if } z = 1, \end{cases}$$

where $Z = \{0, 1\}$. Explicit calculation shows that

$$\begin{aligned} D_f(\eta^\lambda \parallel \zeta^\lambda) &= \sum_{a \in A, \theta \in \Theta, z \in \{0,1\}} \zeta^\lambda(a, \theta, z) f\left(\frac{\eta^\lambda(a, \theta, z)}{\zeta^\lambda(a, \theta, z)}\right) \\ &= \sum_{a \in A, \theta \in \Theta} \left((1-\lambda)\phi^0_A(a)\phi^0_\Theta(\theta) f\left(\frac{(1-\lambda)\phi^0(a, \theta)}{(1-\lambda)\phi^0_A(a)\phi^0_\Theta(\theta)}\right) + \lambda\phi^1_A(a)\phi^1_\Theta(\theta) f\left(\frac{\lambda\phi^1(a, \theta)}{\lambda\phi^1_A(a)\phi^1_\Theta(\theta)}\right)\right) \\ &= (1-\lambda) \sum_{a \in A, \theta \in \Theta} \phi^0_A(a)\phi^0_\Theta(\theta) f\left(\frac{\phi^0(a, \theta)}{\phi^0_A(a)\phi^0_\Theta(\theta)}\right) + \lambda \sum_{a \in A, \theta \in \Theta} \phi^1_A(a)\phi^1_\Theta(\theta) f\left(\frac{\phi^1(a, \theta)}{\phi^1_A(a)\phi^1_\Theta(\theta)}\right) \end{aligned}$$

$$= (1-\lambda)D_f(\phi^0 \parallel \phi^0_A \otimes \phi^0_\Theta) + \lambda D_f(\phi^1 \parallel \phi^1_A \otimes \phi^1_\Theta)$$

By the law of total probability,

$$D_f(\eta^\lambda_{A,\Theta} \parallel \zeta^\lambda_{A,\Theta}) = D_f((1-\lambda)\phi^0 + \lambda\phi^1 \parallel (1-\lambda)\phi^0_A \otimes \phi^0_\Theta + \lambda\phi^1_A \otimes \phi^1_\Theta)$$

, where $\eta^\lambda_{A,\Theta}$ and $\zeta^\lambda_{A,\Theta}$ are the marginals of η^λ and ζ^λ on $A \times \Theta$, respectively.

Using the projection kernel that maps $(a, \theta, z) \to (a, \theta)$ with probability 1, Proposition 2 implies that

$$D_f(\eta^\lambda_{A,\Theta} \parallel \zeta^\lambda_{A,\Theta}) \leqslant D_f(\eta^\lambda \parallel \zeta^\lambda)$$

Hence,

$$\begin{aligned} &D_f((1-\lambda)\phi^0 + \lambda\phi^1 \parallel (1-\lambda)\phi^0_A \otimes \phi^0_\Theta + \lambda\phi^1_A \otimes \phi^1_\Theta) \\ &\qquad\qquad \leqslant (1-\lambda)D_f(\phi^0 \parallel \phi^0_A \otimes \phi^0_\Theta) + \lambda D_f(\phi^1 \parallel \phi^1_A \otimes \phi^1_\Theta) \end{aligned}$$

and thus the mixture distribution must have f-information at most ϵ.

□

Theorems 3 and 4 together imply that the correspondence $F_{f,\epsilon,\mu}$ is induced by a set of information structures that is both individual garbling complete and public randomization complete.

Note that individual garbling completeness is a direct consequence of the data processing inequality. Hence, we could replace D_f in the definition of $F_{f,\epsilon,\mu}$ with any functional that satisfies the DPI and still maintain individual garbling completeness of the correspondence. For example, *Renyi divergence* is a well known dissimilarity measure that is not an f-divergence but does satisfy a DPI. Thus, if we defined a feasibility correspondence for the Renyi divergence analogously to $F_{f,\epsilon,\mu}$, it would be individual garbling complete. We therefore have the following generalization of Theorem 4:

Proposition 3. *For any generalized dissimilarity measure D satisfying the data processing inequality, the correspondence*

$$F_\epsilon(A) = \{\phi \in \Delta(A \times \Theta) | D(\phi \parallel \phi_A \otimes \phi_\Theta) \leqslant \epsilon\}$$

is garbling complete and therefore individual garbling complete.

Proof. By garbling complete we mean that $\phi \in F_\epsilon(A) \Rightarrow \phi' \in F_\epsilon(A')$ if there exists $b : A \to \Delta(A')$ such that $\phi'(a', \theta) = \sum_{a \in A} b(a'|a)\phi(a, \theta)$. The result then follows directly from the proof of Theorem 4. □

Note that we did not need to fix a prior here or specify any properties beyond the DPI. Though this version of Theorem 4 may seem more natural and powerful, we focus on f-divergences due to their practicality, canonical place in the information theory literature and convexity. In particular, the convexity of $F_\epsilon(A)$ does not follow from the DPI.

Among f-divergences, the total variation distance, defined by $f(x) = |x - 1|$, is particularly tractable for linear programming purposes because the total variation bounds can be written as a finite number of linear constraints, through the suitable introduction of auxiliary variables.[5] Let $F^{TV}_{\epsilon,\mu}$ be $F_{f,\epsilon,\mu}$ where $f(x) = |x - 1|$. If $\epsilon \geqslant 2$, then $F^{TV}_{\epsilon,\mu}(A)$ is just the outcomes with marginal μ, and $F^{TV}_{\epsilon,\mu}(A) \cap \text{BCE}(G)$ is the set of BCE with prior μ and no upper bound on information. If $\epsilon = 0$, then $F^{TV}_{\epsilon,\mu}(A)$ contains all outcomes in which a and θ are independent, and $F^{TV}_{\epsilon,\mu}(A) \cap \text{BCE}(G)$ are the correlated equilibria of G when the players have no information and the prior is μ. Thus, $F^{TV}_{\epsilon,\mu}(A)$ interpolates smoothly between the cases of unrestricted information and no information (but players still have access to pure correlation devices).

4.3 Application: Coordinated attack

We conclude this section with an application to a coordinated attack problem. Suppose $\Theta = \{-1, 1\}$, $A_i = \{0, 1\}$, and $u_i(a, \theta) = a_i(\theta + ba_j + c)$ for constants $b, c \in \mathbb{R}$. Both states are equally likely: $\mu(-1) = \mu(1) = 1/2$. We first fix $b = -1/2$ and $c = -1/4$. We ask: What is the maximum probability that $a = (1, 1)$ across all BCE, meaning that both players "attack?"

We first consider the optimum without any feasibility constraints. The unconstrained BCE that maximizes the probability of $a = (1, 1)$ has the following form:

$\theta = -1$

a_1/a_2	1	0
1	δ	0
0	0	$1/2 - \delta$

$\theta = 1$

a_1/a_2	1	0
1	1/2	0
0	0	0

Obedience for $a_i = 1$ reduces to $\delta(-7/4) + (1/2)(1/4) \geqslant 0$, which holds if and only if $\delta \leqslant 1/14$. The optimal BCE makes this constraint bind, in which case the total probability

[5] Total variation distance is often defined using $f(x) = \frac{1}{2}|x - 1|$. Throughout the paper we used the scaled version for ease of exposition.

of $a = (1, 1)$ is 4/7. Note that the optimal BCE also has the form of a public bad news signal. We denote this BCE by ϕ_1.

We now consider what happens with feasibility constraints, meaning that $\phi \in F^{TV}_{\epsilon,\mu}$. When $\epsilon = 0$, the players have no information, and $a_i = 0$ strictly dominates. Hence, the unique feasibility-constrained BCE puts probability one on $(0, 0)$. We denote this BCE by ϕ_0.

When $\epsilon \leqslant 6/7$, the total variation-constrained optimal BCE is simply $\phi_1\epsilon/(6/7) + \phi_0(1 - \epsilon/(6/7))$. This satisfies the total variation constraint due to the data processing inequality. Note that the optimal feasibility-constrained BCE still has the form of a binary public signal, and players are indifferent between entering and not entering after good news. Naturally, as ϵ decreases, the feasibility constraints become tighter, and fewer outcomes can be attained in equilibrium. As $\epsilon \to 0$, we converge to the outcome in which $a = (0, 0)$ occurs with probability one.

To make the example a bit richer, let us now allow the players to make a half investment of $a_i = 1/2$, but maintain the same payoff structure. Thus, the payoff from $(1/2, 1/2)$ is $-3/4$ in state $\theta = -1$ and $1/4$ in state $\theta = 1$. By following a similar analysis, we conclude that the unconstrained BCE that maximizes the probability of $(1/2, 1/2)$ involves a public signal, so that $(1/2, 1/2)$ is played with probability one in the good state and with probability ζ in the bad state, so as to satisfy the obedience constraints with equality: $\zeta(-3/2) + 1/2(1/2) = 0 \iff \zeta = 1/6$, so that the total probability of $(1/2, 1/2)$ is 2/3. We denote this BCE by $\phi_{1/2}$. The resulting total variation in this BCE is 2/3, which is strictly less than the total variation of 6/7 for the BCE that maximizes the probability of $a = (1, 1)$. Indeed, because deviations from $a = (1/2, 1/2)$ are less attractive, the players do not need as much information to be incentivized to play these actions.

As a final exercise, consider the optimal BCE when we maximize the probability of $(1, 1)$ plus η times the probability of $(1/2, 1/2)$. As long as $\epsilon \leqslant (4/7)/(2/3) = 6/7$, the unconstrained optimum will be the BCE that maximizes the likelihood of $(1, 1)$. But if η is sufficiently high, then the optimum changes for $\epsilon \leqslant 6/7$. Rather than stochastically receiving the no-information outcome, in order to satisfy the total variation constraint, the optimal BCE instead is a convex combination of $\phi_{1/2}$ and ϕ_1, with appropriate weights to make the total variation constraint bind. In particular, this happens as long as $\eta(2/3) \geqslant (4/7)(2/3)/(6/7) \iff \eta \geqslant 2/3$. When $\epsilon = 2/3$, the optimum is $\phi_{1/2}$, and for $\epsilon \leqslant 2/3$, the optimal BCE is a mixture of $\phi_{1/2}$ and ϕ_0.

This example illustrates several ideas. First, it shows that individual garbling complete sets of information structures provide a tractable methodology for analyzing BCE with upper bounds on information. Second, it shows that as information constraints become tighter, extremal BCE will change so as to place greater weight on actions that can be incentivized with less information, such as by transitioning from playing $a = (1, 1)$ to playing $a = (1/2, 1/2)$. Finally, the example illustrates consequences of the linearity of the data processing inequality, namely that any objective will be concave in the allowed f-information.

4.4 Other f-information bounds

While we have focused on bounds on the players' f-information about the state, the methodology introduced in this section could be easily adapted to express other kinds of feasibility constraints, such as an upper bound on how much one player knows or a group of players know about the state. We revisit restrictions of this form in Section 7.

5 Linear games

5.1 Robust predictions

We now let Θ be a subset of a linear space, which we will typically think of as $\mathbb{R}^K$. (To be consistent with the earlier sections we will only consider outcomes with a finite support on Θ.) Fix a game $G = (A, u)$, where the payoff $u_i : A \times \Theta \to \mathbb{R}$ is a linear function of $\theta \in \Theta$ for each $a \in A$. We call (A, u) satisfying these conditions a *linear game*.

For a $\mu \in \Delta(\Theta)$ with a finite support, with a slight abuse of notation let $\text{BCE}(\mu; G)$ be the set of BCE in G whose marginal over the states is μ. We will abbreviate $\text{BCE}(\mu; G)$ to $\text{BCE}(\mu)$ when the game G is clear from the context.

For an outcome $\phi \in \Delta(A \times \Theta)$, let $\eta_\phi(a)$ be the *interim (expected) state* conditional on a, i.e., $\eta_\phi(a) = \sum_{\theta \in \Theta} \theta \phi(a, \theta)/\phi_A(a)$ when $\phi_A(a) > 0$, and let $\nu(\phi)$ be the distribution of $\eta_\phi(a)$ under ϕ_A.

We are interested in the extreme points of the set $BCE_F(G)$ where F is a divergence constrained feasibility correspondence. In particular, we will calculate the element of $BCE_F(G)$ that minimizes the expectation of some welfare criterion $w : A \times \Theta \to \mathbb{R}$. We further assume that w is linear in θ for each a. Then the expectation of w given an outcome ϕ is

$$W(\phi) = \sum_{a \in A, \theta \in \Theta} w(a, \theta) \phi(a, \theta).$$

We write $W(\phi; G)$ when we want to emphasize the underlying game G.

Our main result in this section is:

Theorem 5. *Fix a linear game $G = (A, u)$ and a prior $\mu \in \Delta(\Theta)$ with a finite support. Then*

$$\min_{\phi \in F_{f,\epsilon,\mu}(A) \cap \mathrm{BCE}(G)} W(\phi) = \min_{\mu' \in P_\mu} \min_{\phi \in \mathrm{BCE}(\mu')} W(\phi), \tag{5}$$

where

$$P_\mu \equiv \{\nu(\phi) : \phi \in F_{f,\epsilon,\mu}(A)\}. \tag{6}$$

Moreover, there exist optimal solutions ϕ^* and (μ', ϕ') to the left- and right-hand sides of (5)*, respectively, such that $\phi'_A = \phi^*_A$ and $\eta_{\phi^*} = \eta_{\phi'}$.*

Thus for a linear game, the minimization over divergence constrained equilibrium outcomes in (5) is reduced to a simpler and more familiar minimization over equilibrium outcomes with a prior equal to μ', where μ' is endogenously chosen from P_μ and is a mean-preserving contraction of the true prior μ. This result is particularly helpful when the minimizing solution over equilibrium outcomes is well understood for every prior; Theorem 5 implies that such a solution for some μ' will also be a solution to problem (5). We illustrate this for revenue minimization in the common-value first-price auction in Section 5.2.

Regarding the interpretation of the program (5), we note that it can be equivalently written as

$$\min_{\phi \in \mathrm{BCE}(\mu)} W(\phi) + \lambda D_f(\phi \parallel \phi_A \otimes \mu)$$

where $\lambda \geqslant 0$ is the Lagrange multiplier on the divergence constraint. The above problem is exactly the multiplier robust-control problem in Hansen and Sargent (2001) where we take the reference outcome to be the information structure where the players have no information about the state. The interpretation is that the no information scenario is the analyst's best guess for the agents' information structure, but the analyst does not fully trust it. Instead, the analyst considers many other information structures to be plausible, with plausibility diminishing with their divergence from the no information scenario.

In general, the set P_μ in Theorem 5 will not be convex, since for each prior in P_μ the support can be arbitrary but with at most $|A|$ elements. Nonetheless, to compute the welfare guarantee across divergence constrained equilibrium outcomes in Theorem 5, it is without loss to convexify P_μ, as shown in the following proposition (whose proof is in the appendix). This convexity will prove crucial for maxmin mechanism design in Section 6. Given a set X in a linear space, we denote its convex hull by $\operatorname{conv} X$.

Proposition 4. *The value of problem* (5) *is equal to*

$$\min_{\mu' \in \operatorname{conv} P_\mu} \min_{\phi \in \mathrm{BCE}(\mu')} W(\phi). \tag{7}$$

The intuition for Proposition 4 is that while a convex combination of priors from P_μ need not be in P_μ, for any $\mu \in \operatorname{conv} P_\mu$ and outcome ϕ with marginal μ, the associated distribution of interim states is still in P_μ, and only the interim states matter for a linear game.

Note that Theorem 5 does not rely on a fixed prior or specific properties of f-divergences. Rather, it can be extended to any feasibility correspondence that satisfies the DPI. In contrast, Proposition 4 does rely on convexity properties of f-divergences. We therefore have the following generalization of Theorem 5, analogous to Proposition 3.

Proposition 5. *Fix a linear game $G = (A, u)$. Let F_ϵ be a feasibility correspondence as defined in Proposition 3 such that for all $\phi \in F_\epsilon$, ϕ_θ has finite support. Then*

$$\min_{\phi \in F_\epsilon(A) \cap \mathrm{BCE}(G)} W(\phi) = \min_{\mu' \in P} \min_{\phi \in \mathrm{BCE}(\mu')} W(\phi),$$

where

$$P \equiv \{\nu(\phi) : \phi \in F_\epsilon(A)\}. \tag{8}$$

Moreover, there exist optimal solutions ϕ^* and (μ', ϕ') to the left- and right-hand sides of (5)*, respectively, such that $\phi'_A = \phi^*_A$ and $\eta_{\phi^*} = \eta_{\phi'}$.*

Proof. The result follows from the proof of Theorem 5. □

5.2 Application: First-price auctions with common values

We now illustrate our results on linear games with the example of a first-price auction for a common-value good, as in Bergemann, Brooks, and Morris (2017). The state is the common value of the good, which is drawn from a finite subset Θ of $\mathbb{R}$. Each player i's action is a bid, which is an element of finite set $A_i \subset \mathbb{R}$. Payoffs are given by

$$u_i(a, \theta) = \begin{cases} \frac{1}{|H(a)|}(\theta - a_i) & i \in H(a); \\ 0 & \text{otherwise,} \end{cases}$$

$$H(a) = \left\{ j \,\middle|\, a_j = \max_{j'=1,\dots,N} a_{j'} \right\},$$

for $a \in A$ and $\theta \in \Theta$. In other words, the winner of the good is one of the high bidders, breaking ties randomly, and the winner pays their bid. We note that u_i is linear in the state, so this is a linear game. The welfare criterion is revenue, i.e., $w(a, \theta) = \max_{i=1,\ldots,N} a_i$.

For our simulations, we will assume that the prior μ is uniform on $\Theta = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. The action space is an evenly spaced grid of 100 points on $[0, 0.5]$.

We first computed the extreme point of the set of total variation constrained BCE that minimizes expected revenue. We refer to the minimum value as the first-price auction's *revenue guarantee*. The guarantee implicitly depends on the bound on the players' information about the state, in terms of total variation distance. The results of the calculation are depicted in Figures 2 and 3. On the left-hand side of Figure 2, we have plotted the marginal over the action profile, i.e., ϕ_A. On the right-hand side, we have plotted the function η_ϕ. Each of these objects is depicted for values of $\epsilon = \{0.5, 1, 2\}$. While not necessarily self-evident from the figures, it is easily verified that in each case, the players' actions are independent and identically distributed. Moreover, we can see that the interim value $\eta_\phi(a)$ is only a function of the highest of the bidders' signals. As the left-hand side of Figure 3 shows, as ϵ decreases, the distribution of interim values $\nu(\phi)$ becomes more and more compressed, and when $\epsilon = 0$ it is simply a point mass on the prior expectation of 1/2.

The BCE depicted in these figures closely resemble those described by Bergemann, Brooks, and Morris (2017), who studied revenue-minimizing BCE of first-price auctions, in a model with continuous values and continuous bids; their results correspond to the case of $\epsilon = 2$ in our simulation since in this case the total variation constraint does not bind. Bergemann, Brooks, and Morris (2017) show that in the case of a pure common value, revenue is minimized when the bidders receive iid signals, the high signal is equal to the common value, and the bidders use monotonic pure strategies. In fact, the bidders treat their signals as if they were private values, and play the standard Vickrey equilibrium of the independent private value first-price auction. The only substantive difference between our simulations and the structure identified in Bergemann, Brooks, and Morris (2017) is that in the simulations, the high signal (which is one-to-one with the high bid) does not reveal the true value (unless $\epsilon = 2$), but rather reveals a noisy estimate of the value. But this is the revenue-minimizing BCE, as identified in Bergemann, Brooks, and Morris (2017), if we treated the interim expected value as the true value. The simulation is therefore consistent with Theorem 5: for linear games, extremal divergence-constrained BCE are extremal BCE under a contracted prior.

We also computed the revenue guarantee over Kullback-Leibler constrained BCE. Because the divergence is calculated using the function $f(x) = x \log(x)$, the program of minimizing revenue over BCE subject to an upper bound on the Kullback-Leibler divergence

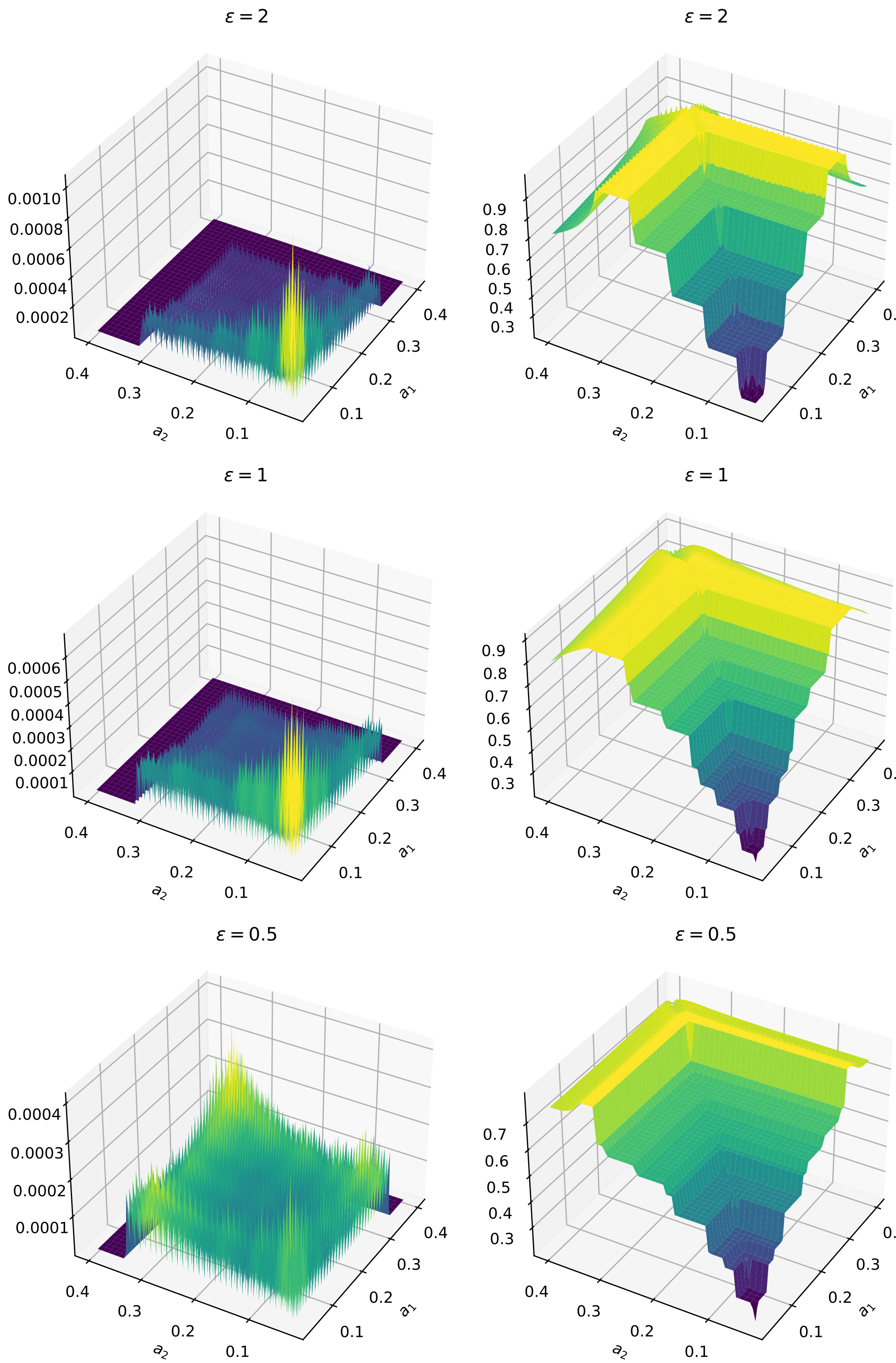

Figure 2: Common value first-price auctions under total variation constraints. On the left-hand side is the marginal joint distribution over action profiles. On the right-hand side is the interim expected value, conditional on the players' actions.

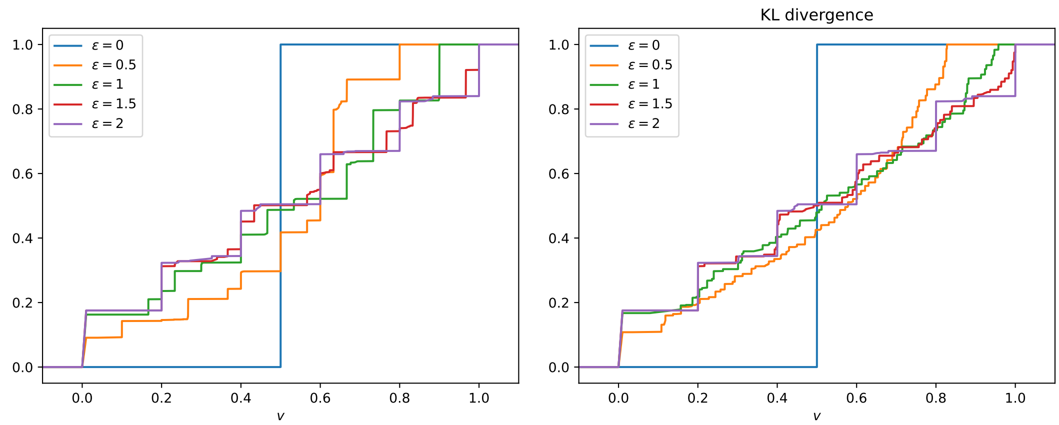

Figure 3: Left: Interim value distributions in the first-price auction under total variation constraints. Right: Interim value distributions in the first-price auction under Kullback-Leibler information constraints.

is not a finite-dimensional linear program. In our simulations, we replace $f(x) = x\log(x)$ by a piecewise linear approximation

$$\tilde{f}(x) = \max_{i=1,\ldots,N} f(x_i) + f'(x_i)(x - x_i) - C, \tag{9}$$

where $f(x_i) + f'(x_i)(x - x_i)$ is a linear approximation of $f(x) = x\log(x)$ around a base point x_i, $f'(x) = 1 + \log(x)$, and C is a constant that ensures $\tilde{f}(1) = 0$. We choose the base points x_i to be an evenly spaced grid of 20 points on $[0, 2]$ as well as $x_i = 4, 6$. (We have $x = \frac{\phi(a,\theta)}{\phi_A(a)\mu(\theta)} \leqslant 6$ since μ is uniformly distributed on 6 values.) In Figure 4 we plot $f(x)$ and $\tilde{f}(x)$ on [0,6] to give a visual sense of the difference, which in a loose sense appears small.

As with the total variation constraint, the extremal Kullback-Leibler constrained BCE are extremal BCE under a contracted prior, which are the interim value distributions plotted in the right-hand side of Figure 3. Compared with the total variation constraint in Figure 3, we see that the interim value distributions from the Kullback-Leibler information constraint tend to be smoother.

6 Maxmin mechanism design

6.1 A general result

As a further application of our results on f-divergence constrained outcomes in linear games, we will consider a variation on the informationally robust mechanism design model

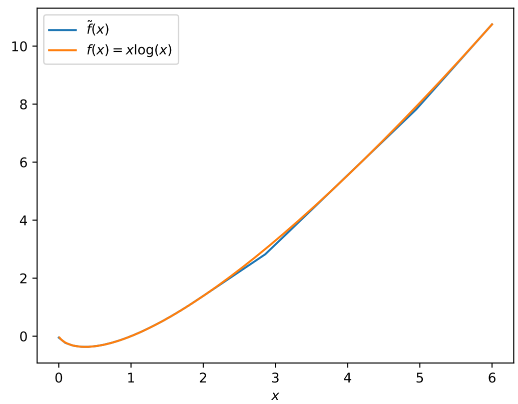

Figure 4: Piecewise linear approximation of $f(x) = x\log(x)$ (cf. equation (9)).

of Brooks and Du (2024b), but where we now posit that the players have limited information about the state. Suppose a mechanism designer wants to maximize the mechanism's performance guarantee over the information structures subject to an upper bound on the agents' information. The designer controls an outcome $\omega \in \Omega$. Suppose the agents' and designer's utilities, $\tilde{u}_i(\omega, \theta)$ and $\tilde{w}(\omega, \theta)$, are linear in the state θ for each fixed ω.

A mechanism is a tuple $M = (A, m)$ where $A = \prod_{i=1,\ldots,N} A_i$ and $m : A \to \Delta(\Omega)$. A mechanism M induces a game where $u_i(a, \theta; M) = \sum_{\omega\in\Omega} m(\omega \mid a)\tilde{u}_i(\omega, \theta)$ and $w(a, \theta; M) = \sum_{\omega\in\Omega} m(\omega \mid a)\tilde{w}(\omega, \theta)$. A mechanism M is *participation secure* if for every player i there exists an action $0 \in A_i$ such that $u_i(0, a_{-i}, \theta; M) \geqslant 0$ for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$. Recall our definitions of $W(\phi; M)$ and $\text{BCE}(\mu; M)$ where we identify the game with its underlying mechanism M.

Following Brooks and Du (2024b), we consider action spaces parameterized by an integer k, where the kth action space has $k^2 + 1$ actions that are labeled $A_i = \{0, 1/k, 2/k, \ldots, k\}$. Let $\mathcal{M}_k^0$ be the set of participation-secure mechanisms on A where 0 is a participation secure action for every player.

For a given $\mu \in \Delta(\Theta)$ and $\epsilon \geqslant 0$, the performance guarantee of a mechanism M over f-divergence constrained equilibrium outcomes is

$$\min_{\phi \in F_{f,\epsilon,\mu}(A)\cap \text{BCE}(M)} W(\phi; M). \tag{10}$$

Thus, the guarantee-maximizing mechanism in $\mathcal{M}_k^0$ solves

$$\max_{M=(A,m)\in\mathcal{M}_k^0} \min_{\phi \in F_{f,\epsilon,\mu}(A)\cap \text{BCE}(M)} W(\phi; M). \tag{11}$$

By Theorem 5 and Proposition 4, problem (11) is equivalent to:

$$\max_{M=(A,m)\in\mathcal{M}_k^0} \min_{\mu'\in\text{conv } P_\mu} \min_{\phi\in\text{BCE}(\mu';M)} W(\phi; M). \tag{12}$$

It is natural to ask how this program is related to

$$\min_{\mu'\in\text{conv } P_\mu} \max_{M\in\mathcal{M}_k^0} \min_{\phi\in\text{BCE}(\mu';M)} W(\phi; M). \tag{13}$$

If these two programs have the same value, then we would know that the solution to (12) reduces to the solution of the analogous maxmin problem with a different prior and no information constraints. This would be especially useful in cases where the solution to the maxmin problem with a fixed prior is known for all priors, such as in the common-value auction model studied in Brooks and Du (2021).

It is immediate that (12) is less than or equal to (13). However, it does not follow from the standard minimax theorem that (12) is equal to (13): for a fixed μ', $\min_{\phi\in\text{BCE}(\mu';M)} W(\phi; M)$ is generally neither a concave nor convex function of M.

On the other hand one can bound the gap between (12) and (13) by applying the bounding programs from Brooks and Du (2024b). In particular, Theorem 1 of Brooks and Du (2024b) shows that (12) is at least

$$\max_{M=(A,m)\in\mathcal{M}_k^0} \min_{\mu'\in\text{conv } P_\mu} \sum_{\theta\in\Theta} \mu'(\theta) \min_{a} \sum_{\omega\in\Omega} \left[\tilde{w}(\omega,\theta) m(\omega|a) + \sum_{i=1,\ldots,N} \tilde{u}_i(\omega,\theta) \nabla_i^+ m(\omega|a) \right], \tag{14}$$

where

$$\nabla_i^+ f(a) = \begin{cases} (k-1)(f(a_i+1/k, a_{-i}) - f(a)) & \text{if } a_i < k; \\ 0 & \text{if } a_i = k. \end{cases}$$

The inner function in (14) is clearly linear in μ' for a fixed m and concave in m for a fixed μ'. Thus, by Sion's minimax theorem, (14) is equal to

$$\min_{\mu'\in\text{conv } P_\mu} \max_{(A,m)\in\mathcal{M}_k^0} \sum_{\theta\in\Theta} \mu'(\theta) \min_{a\in A} \sum_{\omega\in\Omega} \left[\tilde{w}(\omega,\theta) m(\omega|a) + \sum_{i=1,\ldots,N} \tilde{u}_i(\omega,\theta) \nabla_i^+ m(\omega|a) \right]. \tag{15}$$

Moreover, (13) is clearly less than or equal to

$$\min_{\mu'\in\text{conv } P_\mu} \min_{I=(S,\sigma):\sigma_\Theta=\mu'} \max_{M\in\mathcal{M}_k^0} \max_{\phi\in E_I(M)} W(\phi; M),$$

where we again identify M with the induced game structure. Brooks and Du (2024b) refer to the inner double maximand as the *potential* of an information structure I. By Theorem 1 of Brooks and Du (2024b), the minimum potential across all I with prior μ is at most

$$\min_{\mu' \in \operatorname{conv} P_\mu} \min_{\phi \in \Delta(A \times \Theta): \phi_\Theta = \mu'} \sum_{a \in A} \max_{\omega \in \Omega} \sum_{\theta \in \Theta} \left[\tilde{w}(\omega, \theta)\phi(a, \theta) - \sum_{i=1,\ldots,N} \tilde{u}_i(\omega, \theta) \tilde{\nabla}_i^+ \phi(a, \theta) \right], \quad (16)$$

where

$$\tilde{\nabla}_i^+ f(a) = \begin{cases} -f(k, a_{-i}) & \text{if } a_i = k; \\ f(k, a_{-i}) - k f(k - 1/k, a_{-i}) & \text{if } a_i = k - 1/k; \\ k(f(a_i + 1/k, a_{-i}) - f(a)) & \text{otherwise.} \end{cases}$$

Thus, we conclude that the value of (15) is less than that of (12), which is less than that of (13), which is in turn less than (16). This proves the following result:

Theorem 6. *Let* (μ', m) *be a Nash equilibrium of the zero-sum game in* (15)*. Then the guarantee of mechanism* M *over* f*-divergence constrained equilibrium outcomes, given by* (10)*, is at least the value of* (15)*. Moreover, the guarantee of any mechanism is at most the value of problem* (16)*.*

Thus, if the values of (15) and (16) are equal, then mechanisms maximizing (15) must approximately maximize the guarantee. In fact, Brooks and Du (2024a) have shown that the difference between the inner max in (15) and the inner min in (16) must tend to 0 as $k \to \infty$ for each μ'.

6.2 Application: Common value auction design

We simulated solutions of the programs (15) and (16) for the common value auction model studied in Brooks and Du (2021). In particular, Θ is a finite subset of $\mathbb{R}$. The outcome $\omega = (q, t) \in \mathbb{R}_+^N \times \mathbb{R}^N$ is a collection of allocations and transfers, where $\sum_{i=1,\ldots,N} q_i = 1$.[6] Player i's utility is

$$\tilde{u}_i(q, t, \theta) = \theta q_i - t_i.$$

The designer's objective is revenue maximization:

$$\tilde{w}(q, t, \theta) = \sum_{i=1,\ldots,N} t_i.$$

[6]This corresponds to the "must-sell" case studied in Brooks and Du (2021).

Thus, the program (15) is a lower bound on the maximum guarantee for revenue across all mechanisms that always sell the good, and the program (16) is an upper bound on the minimum potential for revenue across all information structures, again assuming that the good must be sold.

We first solved programs (15) and (16) under the total variation information constraints when $k = 15$, meaning that each player has 226 actions, and μ is the uniform distribution on $\Theta = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ (as in Section 5.2).

The results of the simulations are depicted in Figures 6–8. Figure 6 depicts the information structures that minimize the upper bound on the potential (16). On the left-hand side is the joint distribution of the players' signals, and on the right-hand side is the players' expectation of the value conditional on the signal profile. There are two notable features: First, the likelihood of a signal profile only depends on the sum of the signals. While this is again not self-evident from the figure, it is also easily verified that the distribution of signals is independent, and these two properties together imply that the signals are iid exponential random variables. Second, the interim expected value is a non-decreasing function of the sum of the players' signals.

This is the same structure as identified in Brooks and Du (2021) for the potential minimizing information structure with a fixed prior, which coincides with the information in Figure 6 when $\epsilon = 2$.[7] In particular, the signals are iid exponential and the interim value is a non-decreasing function of the sum of the signals. The only difference is in the particular interim value function. Without any constraints on the players' information, the sum of the signals fully "reveals" the value, meaning that the distribution of the interim value is equal to the distribution of the ex post value. However, with an upper bound on information, the sum of the signals is only a noisy signal about the value. The particular noisy interim value distributions, which are mean-preserving contractions of the true prior, are depicted in the left panel of Figure 7. Consistent with Theorems 5 and 6, and according to the limit analysis of Brooks and Du (2021), this "noisy" interim value function would be the potential minimizing value function, without any constraints on the players' information, if we treated the interim value as the ex post value, and replaced the true prior with the corresponding contracted prior in Figure 7.

We obtain analogous results for the mechanisms that maximize the lower bound on the guarantee, depicted in Figure 8. In fact, for each value of ϵ, the computed allocations

[7]As with our discussion of the first-price auction, the analysis of Brooks and Du (2021) is primarily in a limit continuum model, although Section 5 of that paper shows that the potential minimizing information structures can be approached with finite information structures that have the same qualitative features as in the simulations depicted in Figure 6. A similar comment applies to the mechanisms depicted in Figure 8.

look nearly identical, and are close to the proportional allocation $q_i(a) = a_i/\Sigma a$, where $\Sigma a = a_1 + \ldots + a_N$. Consistent with the results of Brooks and Du (2021), there seem to be many optimal transfer rules. In our simulations, we selected a particular transfer rule by imposing an additional constraint, that the aggregate transfer $\sum_{i=1,\ldots,N} t_i(a)$ only depends on the sum of the actions Σa. Adding this constraint did not change the optimal value, and it resulted in transfer rules of the form $t_i(a) = T(\Sigma a)a_i/\Sigma a$, where T depends on the contracted prior and is plotted in the right panel of Figure 7. When $\epsilon = 2$, the aggregate transfer coincides with that in Brooks and Du (2021). When $\epsilon = 0$, the aggregate transfer is close to 0.5 for any positive aggregate action; the aggregate transfer is not exactly 0.5 since we use a discrete approximation of $k = 15$; we can see in Figure 5 that due to the discrete approximation, the revenue guarantee is close to but strictly less than 0.5 when $\epsilon = 0$, even though in this case the designer knows that all agents have no information about the common value. In sum, the mechanisms from the simulations have the form of the proportional auctions described in Brooks and Du (2021) and calibrated to the contracted priors of Figure 7.

The optimal values of programs (15) and (16) are plotted in Figure 5, along with that of the first-price auctions from Section 5.2. We see that the upper and lower bounds are close, and the lower bound is significantly larger than the revenue guarantee from the first-price auction when $\epsilon > 0$. Thus, the proportional auctions in Figure 8 significantly outperform first-price auction in terms of the revenue guarantee.

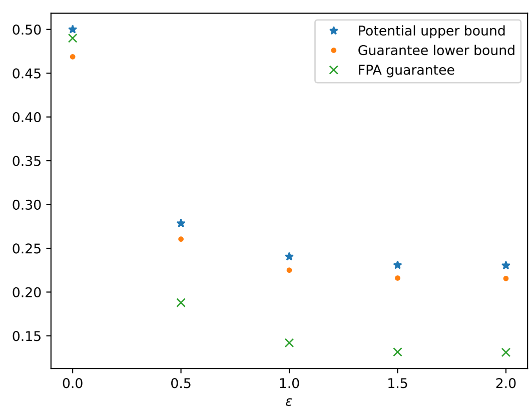

Figure 5: Revenue guarantees and potentials from maxmin common value auctions and first-price auctions under total variation information constraints.

Finally, we depict the corresponding pictures under the Kullback-Leibler information constraints in Figure 9, with the same piecewise linear approximation as in Section 5.2. We see that the interim value distributions, aggregate transfers, and revenues are qualitatively similar to those under total variation information constraints, though just as with the first-

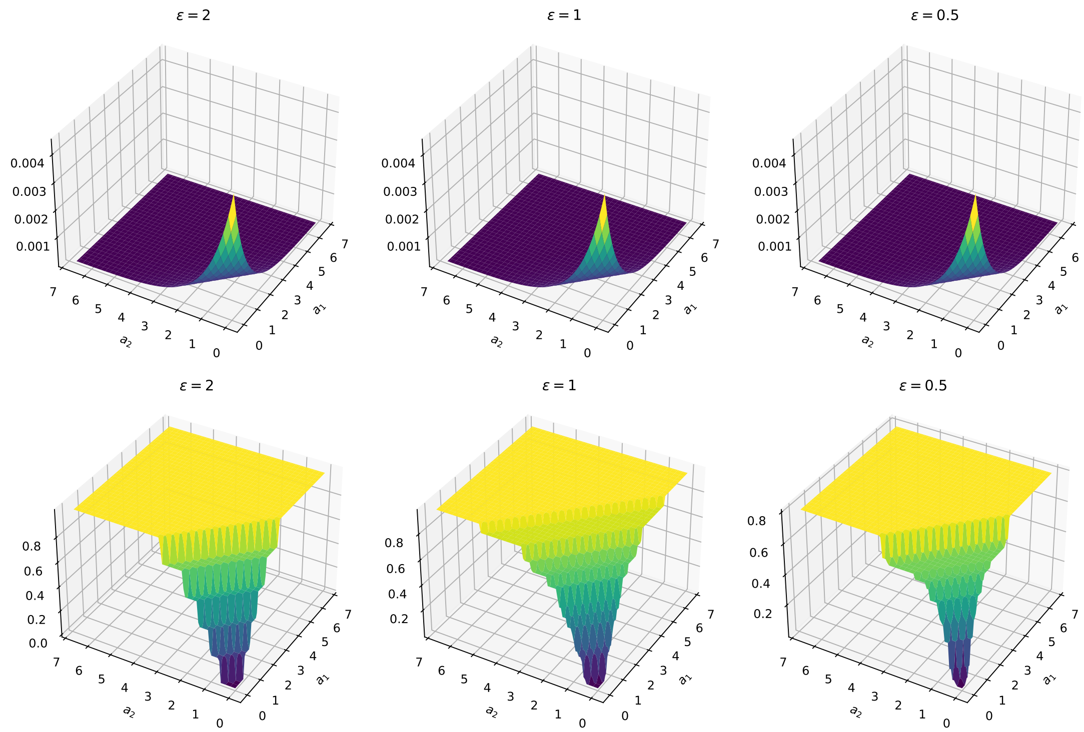

Figure 6: Information structures from program (16) under total variation information constraints; the left is the marginal distribution of the signal profiles, and the right is the interim value as a function of the signal profile.

price auction, the interim value distributions tend to be smoother with Kullback-Leibler information constraints than with total variation distance.

The take-away from this application is that it is possible to combine the insights of this paper with the informationally robust mechanism design approach of Brooks and Du (2024b), in order to obtain new insights about informationally robust mechanisms when the players have bounded information about the state. This approach is especially powerful when the agents' preferences over the mechanism's outcome are linear in the state.

7 Adding a Lower Bound on Information

In this section, we discuss the addition of a lower bound on information into the analysis. We conjecture that all of our results so far would generalize to this case, and we will sketch the arguments, but we do not provide formal proofs. We also discuss generalizations of the divergence constrained BCE from Section 4 that interact with the lower bound, and we apply those constraints to study bidding behavior in private-value first-price auctions.

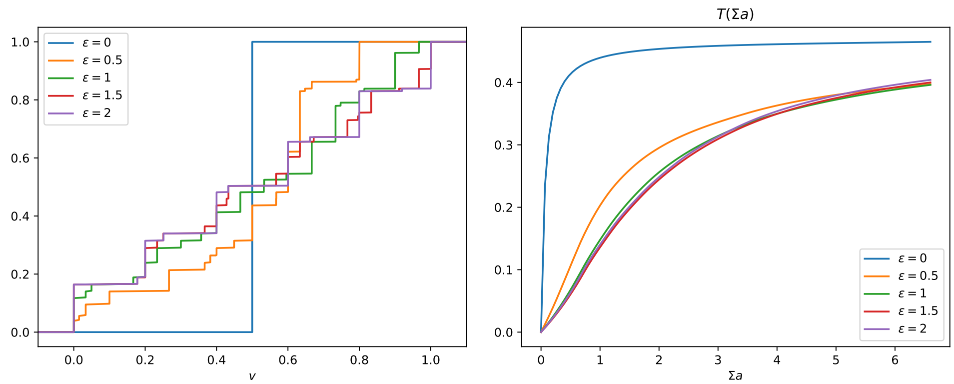

Figure 7: The interim value distributions and aggregate transfers from maxmin common value auctions under total variation information constraints.

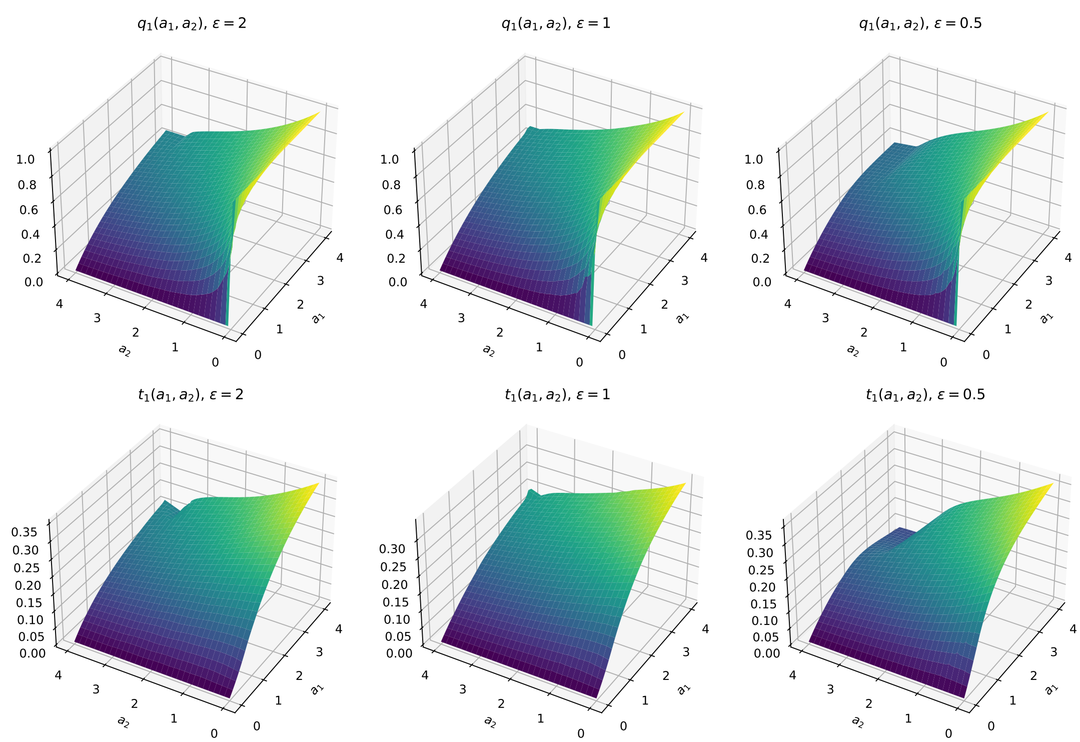

Figure 8: Maxmin common value auctions under total variation information constraints.

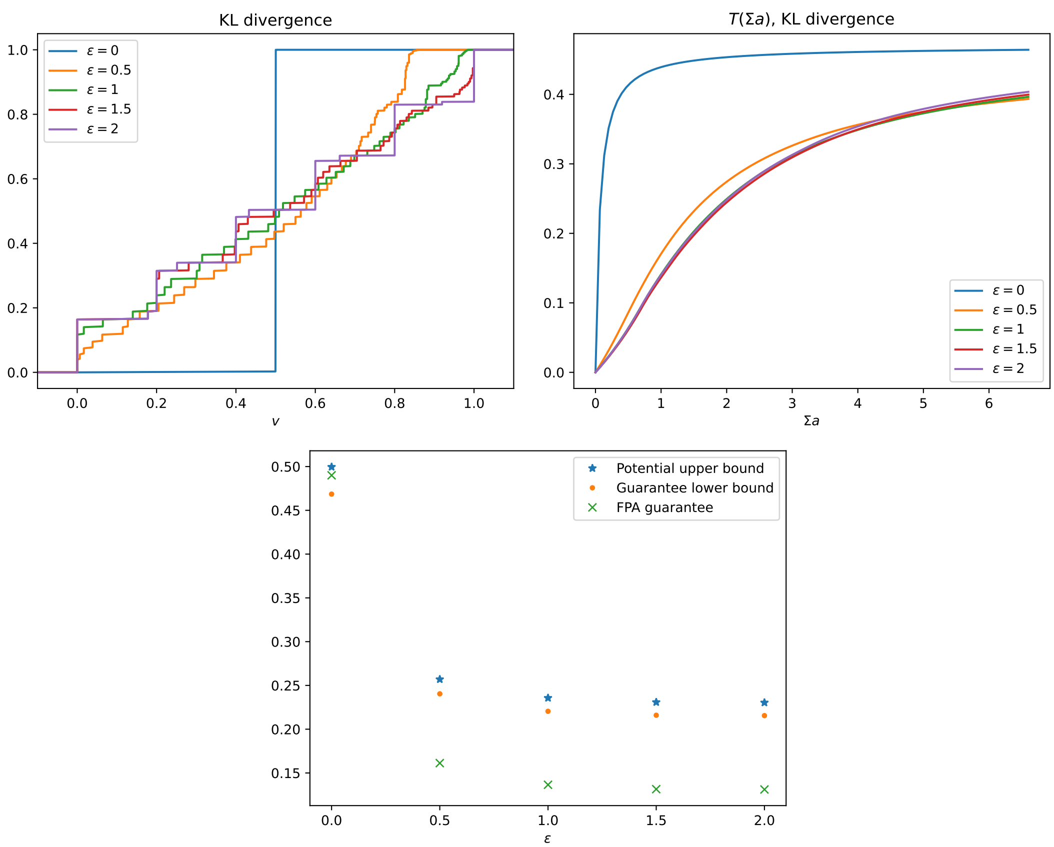

Figure 9: The interim value distributions, aggregate transfers, and revenues from maxmin common value auctions under Kullback-Leibler information constraints.

7.1 Generalizing our results

As mentioned in the introduction, the definition of BCE given in Bergemann and Morris (2016) incorporates a lower bound on information in the following manner: There is a *base information structure* $\underline{I} = (\underline{S}, \underline{\sigma})$. A BCE of a game $G = (A, u)$ is defined to be a joint distribution $\phi \in \Delta(A \times \underline{S} \times \Theta)$ such that the marginal of ϕ on $\underline{S} \times \Theta$ is $\underline{\sigma}$, and such that the following obedience constraints are satisfied: For all i, $\underline{s}_i$, a_i, and a_i',

$$\sum_{a_{-i} \in A_{-i}, \underline{s}_{-i} \in \underline{S}_{-i}, \theta \in \Theta} \phi(a_i, a_{-i}, \underline{s}_i, \underline{s}_{-i}, \theta)(u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta)) \geqslant 0. \tag{17}$$

In other words, conditional on $(a_i, \underline{s}_i)$, a_i maximizes player i's payoff. Thus, the lower bound strengthens the obedience constraint by allowing players to condition their deviation on their base signal $\underline{s}_i$, in addition to their recommended action a_i (as was the case in (1)). Notice that there is also a feasibility constraint on the marginal on $\underline{S} \times \Theta$ (analogous to fixing the marginal on Θ, as we did in Sections 4–6), but there are no other feasibility constraints. Let the set of BCE be denoted by $\text{BCE}(G, \underline{I})$.

The main theorem in Bergemann and Morris (2016) shows that these BCE are precisely the equilibrium outcomes ranging over all I that are more informative than $\underline{I}$ in the sense of *individual sufficiency*, meaning that I is equivalent an information structure of the form $(\prod_{i=1,\ldots,N}(S_i \times \underline{S}_i), \sigma)$, and the marginal of σ on $\underline{S} \times \Theta$ is $\underline{\sigma}$. We will also write $S \times \underline{S}$ for the signal profile space. Informally, I is equivalent to players observing their signals in $\underline{I}$, plus additional signals s in S, which may be correlated with $(\underline{s}, \theta)$ in an arbitrary manner.

The information lower bound could be used in various ways, such as to model the hypothesis that values are private. In particular, we could suppose that under $\underline{I}$, $\underline{s}_i$ reveals to player i all aspects of θ that are payoff relevant to them. It could also be used to model interdependent values using "payoff types", as in Bergemann and Morris (2005), where $\theta = \underline{s}$.

We now explain how the lower bound could be incorporated into our theory. The lower bound $\underline{I}$ can be viewed as consisting of two pieces: One is the assumption that signals have a product form $S \times \underline{S}$ for a fixed set of base signals $\underline{S}$. The second is a feasibility restriction on the marginal on $\underline{S} \times \Theta$. The product signals can be incorporated into our analysis by first defining outcomes as distributions in $A \times \underline{S} \times \Theta$. In the definition of BCE, we would use the stronger obedience constraint (17) in lieu of (1). Also, the set of information structures $\mathcal{I}$ should only consist of signals of the product form. The notion of an equilibrium and strategy profile "inducing" an outcome should be adapted in the obvious way.

The notions of individual garbling and coordinated individual garbling must also be adapted. In particular, $I = (S \times \underline{S}, \sigma)$ is an individual garbling of $I' = (S' \times \underline{S}, \sigma)$ if there are mappings $b_i : S'_i \times \underline{S}_i \to \Delta(S_i)$ such that for all $(s, \underline{s}, \theta)$

$$\sigma(s, \underline{s}, \theta) = \sum_{s' \in S'} b(s|s', \underline{s})\sigma'(s', \underline{s}, \theta),$$

where $b(s|s', \underline{s}) = \prod_{i=1,\ldots,N} b_i(s_i|s'_i, \underline{s}_i)$. Implicit in this definition are the ideas that a player's garbling can depend on their base signal, and the base signal is unchanged by the garbling. The individual garbling is coordinated if the belief about $(s_{-i}, \underline{s}_{-i}, \theta)$ does not depend on s'_i, conditional on $(s_i, \underline{s}_i)$. Formally, if we let $\sigma(s_{-i}, \underline{s}_{-i}, \theta|s_i, \underline{s}_i)$ denote the belief of agent i conditional on $(s_i, \underline{s}_i)$, then a coordinated individual garbling must further satisfy, for all i, $(s'_i, \underline{s}_i)$, and s_i such that $b_i(s_i|s'_i, \underline{s}_i) > 0$,

$$\sigma(s_{-i}, \underline{s}_{-i}, \theta|s_i, \underline{s}_i) = \sum_{s'_{-i} \in S'_{-i}} \prod_{j \neq i} b_j(s_j|s'_j, \underline{s}_j)\sigma'(s'_{-i}, \underline{s}_{-i}, \theta|s'_i, \underline{s}_i).$$

A feasibility correspondence is now a mapping that associates to each product set of action profiles A a set $F(A) \subseteq \Delta(A \times \underline{S} \times \Theta)$. The definitions of individual garbling completeness of a set of information structures and of a feasibility correspondence apply in this more general setting without modification.

With these adjustments, Theorem 1 would remain true as stated. Generalizing the proof of the only if direction is straightforward: For any (I, G) and equilibrium outcome ϕ, there is an associated revelation information structure $I' = (A \times \underline{S}, \phi)$ and equilibrium outcome. If $\mathcal{I}$ is individual garbling complete and $I \in \mathcal{I}$, then I' is a coordinated individual garbling of some $I'' \in \mathcal{I}$. The same argument then shows that there is an equilibrium of (I'', G) that also induces ϕ.

The proof of the if direction of Theorem 1 can be similarly adapted. However, the construction of the separation game must be adjusted. In particular, in the separation game G for an information structure $I = (S \times \underline{S}, \sigma)$, agents are either reporting signals $(s_i, \underline{s}_i)$ or taking the spoiler actions. Under the "revelation" outcome for the separation game, the reported $(s_i, \underline{s}_i)$ is such that the component $\underline{s}_i$ matches its true value. If this revelation outcome is also an equilibrium outcome of G for some information structure $I' = (S' \times \underline{S}, \sigma')$, then for any type $(s'_i, \underline{s}'_i)$ reporting $(s_i, \underline{s}_i)$, it must be that $\underline{s}'_i = \underline{s}_i$. Because this report is preferred by $(s'_i, \underline{s}_i)$ to any of the spoiler actions, the belief conditional on $(s'_i, \underline{s}_i)$ about $(s_{-i}, \underline{s}_i)$ must be the same as that of belief of $(s_i, \underline{s}_i)$ under I, thus proving that I is a coordinated individual garbling of I' (in the modified sense given above). The rest of the proof of the if direction goes through without modification. We conjecture that Theorems 2 and 3 would similarly go through, but we will not sketch the argument.

Most of the substance of the lower bound comes from additional feasibility restrictions on the marginal on $\underline{S} \times \Theta$. As stated above, in Bergemann and Morris (2016), it is assumed that this marginal is the same for all information structures in $\mathcal{I}$, equal to a fixed $\underline{\sigma}$. We could similarly impose this kind of restriction, just as we previously fixed the marginal on Θ.

The bottom line is that with some additional notation, our key results readily generalize to the case where there is a lower bound on information, given by a base information structure.

7.2 Divergence constrained outcomes and other solution concepts

The literature on incomplete information correlated equilibrium studies various feasibility restrictions on the BCE outcomes based on what the players know beyond their base signals. In this section we use our divergence constraints to relax and interpolate these feasibility restrictions.

Fix a base information structure $\underline{I} = (\underline{S}, \underline{\sigma})$. Two prominent incomplete information correlated equilibrium concepts are:

1. *Bayesian solution*: a BCE outcome $\phi \in \Delta(A \times \underline{S} \times \Theta)$ is a Bayesian solution (Forges, 1993) if $\phi(a \mid \underline{s}, \theta)$ is constant over θ for all a, $\underline{s}$. That is, conditional on $\underline{s}$, a and θ are independent. Thus, in an information structure whose equilibrium outcomes are Bayesian solutions, the players do not have information about the state θ beyond what is in the base signal profile.

2. *Belief invariant BCE*: a BCE outcome $\phi \in \Delta(A \times \underline{S} \times \Theta)$ is a belief invariant BCE (Liu, 2015) if and only if
$$\phi(\{a_i\} \times A_{-i} \mid \underline{s}_i, \underline{s}_{-i}, \theta)$$
is constant over $\underline{s}_{-i}$ and θ for all i, a_i, $\underline{s}_i$. That is, conditional on $\underline{s}_i$, a_i and $(\underline{s}_{-i}, \theta)$ are independent. Liu (2015) shows that the belief invariant BCE's are precisely the equilibrium outcomes from information structures that induce the same distribution over hierarchies of beliefs about θ as $\underline{I}$. Thus, in an information structure whose equilibrium outcomes are belief-invariant BCE's, each player may have "redundant" additional information that does not differ from their base signal in terms of the hierarchy of beliefs about θ.

Generalizing these two restrictions, we propose a novel solution concept of *group belief invariant BCE*, which is defined to be a BCE outcome $\phi \in \Delta(A \times \underline{S} \times \Theta)$ such that

$$\phi(\{a_J\} \times A_{-J} \mid \underline{s}_J, \underline{s}_{-J}, \theta)$$

is constant over $\underline{s}_{-J}$ and θ for all $J \subseteq \{1, \ldots, N\}$, a_J, $\underline{s}_J$, where $a_J = (a_i)_{i \in J}$, and likewise for a_{-J}, $\underline{s}_J$ and $\underline{s}_{-J}$. That is, conditional on $\underline{s}_J$, a_J is independent of $(\underline{s}_{-J}, \theta)$. (This condition reduces to the Bayesian solution if $J = \{1, \ldots, N\}$ and to the belief invariant BCE if $J = \{j\}$ for $j = 1, \ldots, N$.) We have the following existence result, which is a corollary of Proposition 8 that we state and prove later in this section:

Proposition 6. *For every $(G, \underline{I})$, a group invariant BCE exists.*

We can relax the restrictions imposed by Bayesian solution and (group) belief invariant BCE using divergence bounds, with the following feasibility correspondence: for $J \subseteq \{1, 2, \ldots, N\}$, define

$$G_{f,\epsilon,\underline{\sigma},J}(A) = \{\phi \in \Delta(A \times \underline{S} \times \Theta) : \phi_{\underline{S} \times \Theta} = \underline{\sigma}, \\ D_f(\phi(\cdot \mid \underline{s}_J)_{A_J \times \underline{S}_{-J} \times \Theta} \parallel \phi(\cdot \mid \underline{s}_J)_{A_J} \otimes \phi(\cdot \mid \underline{s}_J)_{\underline{S}_{-J} \times \Theta}) \leqslant \epsilon, \forall \underline{s}_J\},$$

where we only consider $\underline{s}_J$ that have positive probability under $\underline{\sigma}$.

Proposition 7. *As we increase ϵ from 0 to ∞,*

1. *$G_{f,\epsilon,\underline{\sigma},\{1,2,\ldots,N\}}(A) \cap \mathrm{BCE}(G, \underline{I})$ interpolates from the set of Bayesian solutions (when $\epsilon = 0$) to the set of BCE (when $\epsilon = \infty$).*

2. *$\bigcap_{i=1}^{N} G_{f,\epsilon,\underline{\sigma},\{i\}}(A) \cap \mathrm{BCE}(G, \underline{I})$ interpolates from the set of belief invariant BCE (when $\epsilon = 0$) to the set of BCE (when $\epsilon = \infty$).*

3. *$\bigcap_{J \subseteq \{1,2,\ldots,N\}} G_{f,\epsilon,\underline{\sigma},J}(A) \cap \mathrm{BCE}(G, \underline{I})$ interpolates from the set of group belief invariant BCE (when $\epsilon = 0$) to the set of BCE (when $\epsilon = \infty$).*

As a final topic for this section, we relate our analysis with two other prominent examples of incomplete information correlated equilibrium: correlated equilibrium in the strategic normal form (Cotter, 1991) and correlated equilibrium in the agent normal form (Forges, 1993). Strategic normal form correlated equilibrium refines of BCE with additional obedience constraints: Effectively, when type $\underline{s}_i$ is evaluating deviations, they may condition on the entire *strategy* of player i, which includes recommendations of all types

$\underline{s}_i' \neq \underline{s}_i$. For that reason, it is not within the class of refinements we consider, which only add constraints on feasibility.

Next, consider the agent normal form of G and $\underline{I}$: each positive probability base signal $\underline{s}_i$ is a separate player, with payoff:

$$\sum_{\theta \in \Theta, \underline{s}_{-i} \in \underline{S}_{-i}} u_i(b_i(\underline{s}_i), b_{-i}(\underline{s}_{-i}), \theta) \underline{\sigma}(\underline{s}_{-i}, \theta \mid \underline{s}_i),$$

where each player $\underline{s}_j$ chooses action $b_j(\underline{s}_j)$, and $b_{-i}(\underline{s}_{-i}) = (b_j(\underline{s}_j))_{j \neq i}$. Since $\underline{I}$ is finite, a correlated equilibrium exists in this finite game. Moreover, we can convert its correlated equilibrium $\zeta \in \Delta(\prod_i A_i^{\underline{S}_i})$ to a BCE outcome in $\Delta(A \times \underline{S} \times \Theta)$ by first drawing $(\theta, \underline{s})$ according to $\underline{\sigma}$, and then drawing a according to the marginal distribution of ζ for the "realized" players $\underline{s}_i$, $i = 1, 2, \ldots, N$. In particular, given $\zeta \in \Delta(\prod_i A_i^{\underline{S}_i})$, if we let $\zeta_{\underline{s}}(a)$ be the marginal probability that the players $\underline{s}$ in the agent normal form play $a \in A$, then the outcome induced by ζ is defined by $\phi(a, \underline{s}, \theta) = \underline{\sigma}(\underline{s}, \theta)\zeta_{\underline{s}}(a)$.

Note that by construction, for any set of players I, the distribution of a_I conditional on $(\underline{s}_I, \underline{s}_{-I}, \theta)$ does not depend on $(\underline{s}_{-I}, \theta)$. Hence, if ϕ is induced by an agent normal form correlated equilibrium, then ϕ is group belief invariant. Moreover, since the agent normal form correlated equilibrium is obedient, ϕ must satisfy the obedience constraints as well. We have therefore proven the following:

Proposition 8. *If an outcome ϕ is induced by a correlated equilibrium of the agent normal form, then ϕ is a group belief invariant BCE.*

Combining this result with the fact that a correlated equilibrium exists in the agent-normal form, we conclude that there exists a group belief invariant equilibrium, thus completing the proof of Proposition 6.

How about the converse of Proposition 8? Is every group belief invariant BCE induced by some agent normal form correlated equilibrium? The answer is no, as the following example shows. Let $N = 2$, $A_i = \underline{S}_i = \{0, 1\}$, and $|\Theta| = 1$. Suppose that all base type profiles are equally likely. Moreover, for every $\underline{s}$ except for $(0, 0)$, we have that $\phi((1, 1), \underline{s}, \theta) = \phi((0, 0), \underline{s}, \theta) = 1/8$, meaning that both players take the same action with probability one. At the same time, $\phi((0, 1), (0, 0), \theta) = \phi((1, 0), (0, 0), \theta)$, so when $\underline{s} = (0, 0)$, the players take different actions with probability one. We claim that there is no joint distribution over actions in the agent normal form that could induce this outcome. For if there were, we would have that the types $\underline{s}_1 = 1$ and $\underline{s}_2 = 1$ take the same action with probability one. Moreover, the same is true for types $\underline{s}_1 = 1$ and $\underline{s}_2 = 0$, and also for $\underline{s}_1 = 0$ and $\underline{s}_2 = 1$. But this implies that $\underline{s}_1 = 0$ and $\underline{s}_2 = 0$ take the same action with probability one as well! Thus, agent normal form correlated equilibrium embodies extra feasibility constraints, beyond group belief invariance.

7.3 Application: First-Price Auction with private values

As an illustration, we apply the results of this section to a first-price auction with private values. For the example, there are two buyers who each have values $\theta_i \in \{0, 1\}$, independently distributed and both equally likely. Payoffs are as given in Section 5.2. We further assume that each player bids weakly less than their value. In the base information structure, each player receives a signal which is equal to their own value.

Equilibrium outcomes of this game under different informational assumptions have been studied by Maskin and Riley (2003), Fang and Morris (2006), Āzacis and Vida (2015), and Bergemann, Brooks, and Morris (2013, 2021). The last two papers in particular solve for a robust prediction across all information structures in which players know their values (the latter reinterpreting the model as one of Bertrand duopoly).

In this simple setting with binary values, the low type buyer with $\theta_i = 0$ always bids zero and loses, except when the other buyer's value is also zero. So, all equilibria have the same total surplus, which is 3/4. But there is significant heterogeneity across information structures in equilibria in the allocation of surplus between the buyers and the seller, and across buyers.

We now further refine their results by solving for the feasibility constrained BCE, where we bound the f-divergence between each player's action and the other player's value, conditional on the player's own value. Thus, as ϵ ranges from ∞ to 0, we interpolate between all BCE and the group-belief invariant BCE. Note that for this model, the state is equal to the vector of players' base types, and there are only two players, so that the sets of belief-invariant and group-belief-invariant BCE coincide.

We computed the set of feasibility constrained BCE for various values of ϵ. We also fixed a finite grid of 60 bids, and we computed the maximum weighted sum of the buyers' surpluses. The simulated sets of buyer surplus pairs are depicted in Figure 10. We see that as ϵ goes to zero, the set of equilibrium payoffs converges to $(1/4, 1/4)$, which is the Bayes Nash equilibrium payoff under both full information and under no information.

A striking conjecture that is suggested by this simulation is that the unique group-belief invariant Bayes correlated equilibrium payoff is the same as the payoff in the Bayes Nash equilibrium when the players have no information. We leave it to future work to investigate whether or not this conjecture can be analytically verified, and whether it extends beyond two buyers and and binary values.

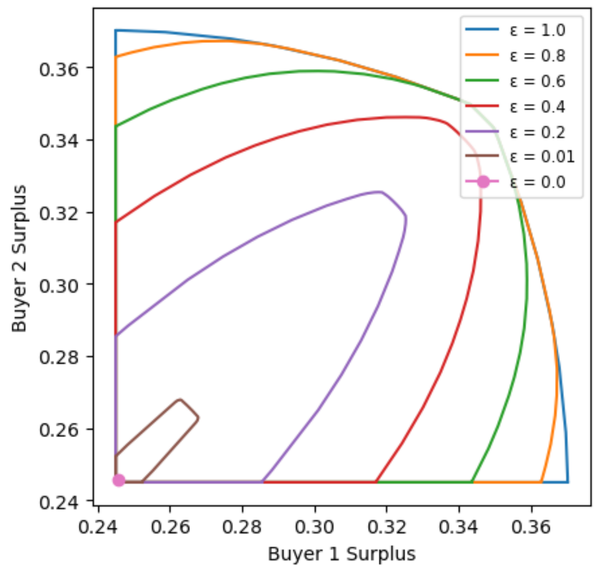

Figure 10: Robust predictions in a private-values first-price auction.

8 Conclusion

The purpose of this paper has been to introduce a new methodology for robust predictions with bounded information. We proposed a new refinement of Bayes correlated equilibria in which we add feasibility constraints. These constraints correspond to a particular class of restrictions on information, namely, those for which the set of admissible information structures is individual garbling complete. Individual garbling completeness precisely characterizes when the restriction on information only constrains which outcomes are feasible, and does not impose additional equilibrium constraints. We have also characterized exactly those feasibility correspondences which can be induced by individual garbling complete sets of information structures, which are those correspondences that satisfy an analogous notion of individual garbling completeness. We have further given an epistemic characterization of when the induced feasibility correspondence is convex, namely, public randomization completeness. We also showed that a feasibility correspondence consisting of those outcomes with a given marginal on θ and an upper bound on the f-divergence between action profiles and states is both individual garbling and public randomization complete. We applied this methodology to a class of linear models, and we determined that extremal BCE with feasibility restrictions correspond to BCE with a set of contracted priors. This finding was illustrated with simulations of extremal BCE of the first-price auction and maxmin auctions.

In future work, we hope to further develop this methodology in several directions. First, we hope to identify additional classes of individual garbling complete feasibility correspondences, analogous to those constrained by f-divergences, that would allow us to more flexibly restrict the amount of information held by each agent. We also hope to incorporate more flexible lower bounds on information into the theory. Finally, we hope to apply our methodology more broadly, to understand better how constraints on information impact informationally robust predictions in settings of economic interest.

References

Āzacis, H. and P. Vida (2015): "Collusive communication schemes in a first-price auction," *Economic Theory*, 58, 125–160.

Bergemann, D., B. Brooks, and S. Morris (2017): "First-Price Auctions with General Information Structures: Implications for Bidding and Revenue," *Econometrica*, 85, 107–143.

——— (2021): "Search, information, and prices," *Journal of Political Economy*, 129, 2275–2319.

——— (2022): "Counterfactuals with Latent Information," *The American Economic Review*, 112.

Bergemann, D., B. A. Brooks, and S. Morris (2013): "Extremal Information Structures in the First Price Auction," *Princeton University William S. Dietrich II Economic Theory Center Research Paper.*

Bergemann, D. and S. Morris (2005): "Robust mechanism design," *Econometrica*, 1771–1813.

——— (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

——— (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

Brooks, B. and S. Du (2021): "Optimal auction design with common values: An informationally robust approach," *Econometrica*, 89, 1313–1360.

——— (2024a): "Dual Reductions and the First-Order Approach for Informationally Robust Mechanism Design," *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

——— (2024b): "On the structure of informationally robust optimal mechanisms," *Econometrica*, 92, 1391–1438.

Cotter, K. D. (1991): "Correlated equilibrium in games with type-dependent strategies," *Journal of Economic Theory*, 54, 48–68.

Fang, H. and S. Morris (2006): "Multidimensional private value auctions," *Journal of Economic Theory*, 126, 1–30.

Forges, F. (1993): "Five legitimate definitions of correlated equilibrium in games with incomplete information," *Theory and decision*, 35, 277–310.

Gossner, O. (2000): "Comparison of information structures," *Games and Economic Behavior*, 30, 44–63.

Hansen, L. and T. J. Sargent (2001): "Robust control and model uncertainty," *American Economic Review*, 91, 60–66.

Lehrer, E., D. Rosenberg, and E. Shmaya (2013): "Garbling of signals and outcome equivalence," *Games and Economic Behavior*, 81, 179–191.

Liu, Q. (2009): "On redundant types and Bayesian formulation of incomplete information," *Journal of Economic Theory*, 144, 2115–2145.

——— (2015): "Correlation and common priors in games with incomplete information," *Journal of Economic Theory*, 157, 49–75.

Maskin, E. and J. Riley (2003): "Uniqueness of equilibrium in sealed high-bid auctions," *Games and Economic Behavior*, 45, 395–409.

Polyanski, Y. and Y. Wu (2023): *Information Theory: From Coding to Learning*, Cambridge University Press (Forthcoming).

A Omitted proofs

Proof of Lemma 1. The action space for each player i is defined as follows. Let D_i^0 be a basis for $\mathbb{R}^{S_{-i}\times\Theta}$, and let $D_i = \{\pm d | d \in D_i^0\}$. Then player i's action set is $A_i = S_i \cup S_i \times D_i$. In other words, the action consists of either a reported signal in I, or a reported signal and direction.

Notice that the "obedient" strategies in which each player reports their true signal would induce the outcome $\phi \in \Delta(A \times \Theta)$, where $\phi(s,\theta) = \sigma(s,\theta)$ for all $(s,\theta) \in S \times \Theta$. For each i and s_i, we define the conditional belief

$$\sigma(s_{-i},\theta \mid s_i) = \sigma(s_i, s_{-i}, \theta) / \sum_{s'_{-i}\in S_{-i},\theta'\in\Theta} \sigma(s_i, s'_{-i}, \theta').$$

Let Ψ_i be the set of the interim beliefs in $\Delta(S_{-i}\times\Theta)$ of the form $\sigma(\cdot \mid s_i)$. We denumerate the elements of $\Psi_i = \{\psi_i^1, \ldots, \psi_i^K\}$, so that for every $k = 1, \ldots, K$, ψ_i^l for $l > k$ are not in the convex hull of $\{\psi_i^1, \ldots, \psi_i^{k-1}\}$.[8] Further let S_i^k be the set of s_i's for which $\sigma(\cdot \mid s_i) = \psi_i^k$.

We will construct players' utilities so that (i) ϕ is an equilibrium outcome of (I, G) and (ii) ψ_i^k is the unique belief in $\Delta(S_{-i}\times\Theta)$ at which s_i is a best response.

We note that whether or not properties (i) and (ii) hold depends only on how we define player i's utilities at action profiles of the form (s,θ) and $((s_i,d), s_{-i}, \theta)$, since these are the only action profiles that can be reached via a single player's deviation from the outcome ϕ.

Now, we inductively define the utilities on $S\times\Theta$. All of the actions in S_i^k have the same utility, which is denoted $u_i^k(s_{-i},\theta)$. Set $u_i^1(s_{-i},\theta) = 0$ for all (s_{-i},θ). Suppose that u_i^l has been defined for all $l < k$. Let $\nu \in \mathbb{R}^{S_{-i}\times\Theta}$ be a separating hyperplane such that

$$\sum_{s_{-i}\in S_{-i},\theta\in\Theta} \nu(s_{-i},\theta)(\psi_i^l(s_{-i},\theta) - \psi_i^k(s_{-i},\theta)) < 0$$

for all $l < k$. We set

$$\begin{aligned} u_i^k(s_{-i},\theta) = 1 + \max_{l<k} &\sum_{s'_{-i}\in S_{-i},\theta'\in\Theta} \psi_i^k(s'_{-i},\theta') u_i^l(s'_{-i},\theta') \\ &+ \gamma_i(s_i)\left(\nu(s_{-i},\theta) - \sum_{s'_{-i}\in S_{-i},\theta'\in\Theta} \nu(s'_{-i},\theta')\psi_i^k(s'_{-i},\theta')\right), \end{aligned}$$

[8] Such an order can be defined inductively. Let ψ_i^K be any extreme point of the convex hull of Ψ_i. Now, having inductively defined $\psi_i^{k+1}, \ldots, \psi_i^K$, let ψ_i^k be any extreme point of the convex hull of $\Psi_i \backslash \{\psi_i^{k+1}, \ldots, \psi_i^K\}$.

where $\gamma_i(s_i) > 0$ large enough that for all $l < k$

$$
\begin{aligned}
&\sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^l(s_{-i},\theta)u_i^k(s_{-i},\theta) \\
&= 1 + \max_{l<k} \sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^k(s_{-i},\theta)u_i^l(s_{-i},\theta) + \gamma_i(s_i) \sum_{s_{-i}\in S_{-i},\theta\in\Theta} \nu(s_{-i},\theta)\left(\psi_i^l(s_{-i},\theta) - \psi_i^k(s_{-i},\theta)\right) \\
&< \sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^l(s_{-i},\theta)u_i^l(s_{-i},\theta).
\end{aligned}
$$

(Such a $\gamma_i(s_i)$ exists because the coefficient on γ is strictly negative and there are only finitely many such inequalities.) At the end of this process, we have constructed the utilities inductively so that if $s_i \in S_i^k$, then

$$
\sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^k(s_{-i},\theta)u_i^k(s_{-i},\theta) > \sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^k(s_{-i},\theta)u_i^l(s_{-i},\theta) \tag{18}
$$

for all $l \neq k$.

It remains to construct the utilities for action profiles of the form $((s_i, d), s_{-i}, \theta)$. For $s_i \in S_i^k$ and $d \in D_i$, we set $u_i((s_i, d), s_{-i}, \theta) = u_i^{k,d}(s_{-i}, \theta)$, where

$$
u_i^{k,d}(s_{-i},\theta) = u_i^k(s_{-i},\theta) + \gamma_i(s_i,d)(d_i(s_{-i},\theta) - \sum_{s'_{-i}\in S_{-i},\theta'\in\Theta} \psi_i^k(s'_{-i},\theta')d_i(s'_{-i},\theta')).
$$

for $\gamma_i(s_i, d) > 0$ small enough such that for all $l \neq k$,

$$
\begin{aligned}
&\sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^l(s_{-i},\theta)u_i^{k,d}(,s_{-i},\theta) \\
&= \sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^l(s_{-i},\theta)u_i^k(s_{-i},\theta) + \gamma_i(s_i,d) \sum_{s_{-i}\in S_{-i},\theta\in\Theta} (d_i(s_{-i},\theta)\left(\psi_i^l(s_{-i},\theta) - \psi_i^k(s_{-i},\theta)\right) \\
&< \sum_{s_{-i}\in S_{-i},\theta\in\Theta} \psi_i^l(s_{-i},\theta)u_i^l(s_{-i},\theta).
\end{aligned} \tag{19}
$$

Such a $\gamma_i(s_i, d) > 0$ exists because there are only finitely many such inequalities and because of the strict inequality (18). This completes the specification of utilities for the separation game.

We now prove properties (i) and (ii). For (i), by (18), we know that at the belief ψ_i^k, an action in S_i^k leads to strictly higher expected utility than any action in $S_i \backslash S_i^k$, and by (19), an action in S_i^k leads to a strictly higher expected utility than (s_i, d) for $s_i \notin S_i^k$. Finally, all actions of the form (s_i, d) with $s_i \in S_i^k$ lead to an expected payoff of

$$\begin{aligned}
&\sum_{s_{-i} \in S_{-i}, \theta \in \Theta} \psi_i^k(s_{-i}, \theta) u_i((s_i, d), s_{-i}, \theta) \\
&= \sum_{s_{-i} \in S_{-i}, \theta \in \Theta} \psi_i^k(s_{-i}, \theta) u_i^k(s_{-i}, \theta) + \gamma_i(s_i, d) \sum_{s_{-i} \in S_{-i}, \theta \in \Theta} d(s_{-i}, \theta) \left(\psi_i^k(s_{-i}, \theta) - \psi_i^k(s_{-i}, \theta)\right) \\
&= \sum_{s_{-i} \in S_{-i}, \theta \in \Theta} \psi_i^k(s_{-i}, \theta) u_i^k(s_{-i}, \theta),
\end{aligned}$$

so that player i is indifferent to all (s_i, d) with $s_i \in S_i^k$. We conclude that if others play the obedient strategies, obedience is a best response for player i, and therefore property (i) is satisfied.

For (ii), suppose that the belief $\psi_i \in \Delta(S_{-i} \times \Theta)$ is not equal to ψ_i^k. Then there is a direction $d \in D_i$ such that

$$\sum_{s_{-i} \in S_{-i}, \theta \in \Theta} d(s_{-i}, \theta)(\psi_i(s_{-i}, \theta) - \psi_i^k(s_{-i}, \theta)) > 0.$$

Hence, the action (s_i, d) yields an expected payoff

$$\begin{aligned}
&\sum_{s_{-i} \in S_{-i}, \theta \in \Theta} \psi_i(s_{-i}, \theta) u_i^k(s_{-i}, \theta) + \gamma_i(s_i, d) \sum_{s_{-i} \in S_{-i}, \theta \in \Theta} d(s_{-i}, \theta)(\psi_i(s_{-i}, \theta) - \psi_i^k(s_{-i}, \theta)) \\
&> \sum_{s_{-i} \in S_{-i}, \theta \in \Theta} \psi_i(s_{-i}, \theta) u_i^k(s_{-i}, \theta),
\end{aligned}$$

and hence s_i is not a best response.

By property (i), $\phi \in E_I(G)$. Now suppose that $\phi \in E_{I'}(G)$ for $I' = (S', \sigma')$, and let b be an equilibrium of (I', G) that induce ϕ. Thus, s_i is a best response at any $s_i' \in S_i'$ for which $b_i(s_i|s_i') > 0$. By property (ii), the belief about (s_{-i}, θ) at s_i' must be $\psi_i^k = \sigma(\cdot, \cdot|s_i)$, so that (2) is satisfied. Hence, I is a coordinated individual garbling of I'. □

Proof of Theorem 3. **If:** Suppose that $F = F_{\mathcal{I}}$ where $\mathcal{I}$ is individual garbling complete and public randomization complete. By Theorem 2, F is individual garbling complete, so it only remains to establish convexity. If $\phi, \phi' \in F(A)$, then they are induced by information structures and strategies $(I = (S, \sigma), b)$ and $(I' = (S', \sigma'), b')$, respectively. Now let $\alpha \in [0, 1]$. Because $\mathcal{I}$ is public randomization complete, the mixture $\alpha I + (1 - \alpha) I'$ is a coordinated individual garbling of some $I'' \in \mathcal{I}$. Let b'' denote the garbling itself. Now, we claim that the following strategies on I'' induce $\alpha\phi + (1-\alpha)\phi'$:

$$\widehat{b}_i(a_i|s_i) = \sum_{\widehat{s}_i \in S_i} b_i(a_i|\widehat{s}_i) b''_i(\widehat{s}_i|s_i) + \sum_{\widehat{s}_i \in S'_i} b'_i(a_i|\widehat{s}_i) b''_i(\widehat{s}_i|s_i).$$

Indeed, using the definition of $\alpha I + (1-\alpha)I'$ and the fact that the mixture is induced by (I'', b''), we have that

$$\begin{aligned}
&\sum_{s \in S''} \widehat{b}(a|s)\sigma''(s,\theta) \\
&= \sum_{s \in S''} \sum_{\widehat{s} \in \prod_{i=1,\ldots,N}\left(S_i \sqcup S'_i\right)} \prod_{i=1,\ldots,N} \left(b_i(a_i|\widehat{s}_i)\mathbb{I}_{\widehat{s}_i \in S_i} + b'_i(a_i|\widehat{s}_i)\mathbb{I}_{\widehat{s}_i \in S'_i}\right) b''(\widehat{s}|s)\sigma''(s,\theta) \\
&= \sum_{\widehat{s} \in \prod_{i=1,\ldots,N}\left(S_i \sqcup S'_i\right)} \prod_{i=1,\ldots,N} \left(b_i(a_i|\widehat{s}_i)\mathbb{I}_{\widehat{s}_i \in S_i} + b'_i(a_i|\widehat{s}_i)\mathbb{I}_{\widehat{s}_i \in S'_i}\right) \left(\alpha\sigma(\widehat{s},\theta)\mathbb{I}_{\widehat{s} \in S} + (1-\alpha)\sigma'(\widehat{s},\theta)\mathbb{I}_{\widehat{s} \in S'}\right) \\
&= \alpha \sum_{\widehat{s} \in S} b(a|\widehat{s})\sigma(\widehat{s},\theta) + (1-\alpha) \sum_{\widehat{s} \in S'} b'(a|\widehat{s})\sigma'(\widehat{s},\theta) \\
&= \alpha\phi(a,\theta) + (1-\alpha)\phi'(a,\theta).
\end{aligned}$$

Hence, $\alpha\phi + (1-\alpha)\phi' \in F_{I''}(A) \subseteq F_{\mathcal{I}}(A) = F(A)$, and therefore F is convex valued.

Only if: By Theorem 2 and Lemmas 3 and 4, if F is individual garbling complete, then $\mathcal{I}_F$, the set of direct recommendation information structures associated with F, is individual garbling complete and induces F. We will further show that $\mathcal{I}_F$ is public randomization complete. To that end, fix $I, I' \in \mathcal{I}_F$ and $\alpha \in [0,1]$. We write $I = (A, \phi)$ and $I = (A', \phi')$. Now, let $\widehat{A}$ be any action space for which $|\widehat{A}_i| = |A_i| + |A'_i|$. Thus, for each i, there is a bijection $\zeta_i : A_i \sqcup A'_i \to \widehat{A}_i$. We write $\zeta(a) = (\zeta_1(a_1), \ldots, \zeta_N(a_N))$. Also define $\widehat{\sigma} \in \Delta(\widehat{A} \times \Theta)$ according to

$$\widehat{\sigma}(s,\theta) = \begin{cases} \alpha\phi(s,\theta) & \text{if } s = \zeta(a) \text{ for } a \in A; \\ (1-\alpha)\phi'(s,\theta) & \text{if } s = \zeta(a') \text{ for } a' \in A'; \\ 0 & \text{otherwise.} \end{cases}$$

It is immediate that $\alpha I + (1-\alpha)I'$ is a coordinated individual garbling of $\widehat{I} = (\widehat{A}, \widehat{\sigma})$, where $b_i(a_i|\zeta_i(a_i)) = 1$. It remains to establish that $\widehat{I} \in \mathcal{I}_F$. As F is individual garbling complete, there are elements $\tilde{\sigma}, \tilde{\sigma'} \in F(\widehat{A})$ given by

$$\tilde{\sigma}(s,\theta) = \begin{cases} \phi(s,\theta) & \text{if } s = \zeta(a) \text{ for } a \in A; \\ 0 & \text{otherwise} \end{cases}, \quad \tilde{\sigma}'(s,\theta) = \begin{cases} \phi'(s,\theta) & \text{if } s = \zeta(a') \text{ for } a' \in A'; \\ 0 & \text{otherwise} \end{cases}.$$

By convexity of F, $\alpha\tilde{\sigma} + (1-\alpha)\tilde{\sigma}' = \widehat{\sigma}$ is in $F(\widehat{A})$, so that $\widehat{I} \in \mathcal{I}_F$. Hence, $\mathcal{I}_F$ is public randomization complete, as desired. □

Proof of Theorem 5. First, let ϕ^* be an optimal solution to the left-hand side of problem (5), and let β^* and η^* be its marginal distribution over actions and interim state function, respectively. Define $\phi'(a,\theta) = \beta^*(a)\mathbb{I}_{\theta=\eta^*(a)}$. By the linearity of the game, we have $\phi' \in \text{BCE}(\nu(\phi^*))$ and $W(\phi^*) = W(\phi')$. Therefore, the optimal value of the left-hand side of (5) is greater than or equal to that of the right-hand side.

Let (μ', ϕ') be an optimal solution to the right-hand side of (5) where $\mu' = \nu(\phi)$ for $\phi \in F_{f,\epsilon,\mu}(A)$. It is without loss of generality to assume that $\eta_\phi(a)$ is unique for all a that occur with positive probability under ϕ. We will construct $\widehat{\phi} \in F_{f,\epsilon,\mu}(A)$ such that $\widehat{\phi}_A = \phi'_A$ and $E_{\widehat{\phi}}[\theta|a] = E_{\phi'}[\theta|a]$ for all a. These conditions ensure that under linearity $\widehat{\phi} \in F_{f,\epsilon,\mu}(A) \cap \text{BCE}(G)$ and $W(\widehat{\phi}) = W(\phi')$. Thus, the optimal value of the right-hand side of (5) is greater than or equal to that of the left-hand side completing the proof. Define:

$$\widehat{\phi}(a,\theta) = \sum_{a' \in A} \phi'(a|\eta_\phi(a'))\phi(a',\theta)$$

Then $\widehat{\phi} \in F_{f,\epsilon,\mu}(A)$ is immediate from the DPI using Kernel $b(a,\theta|a',\theta') = \phi'(a|\eta_\phi(a'))\mathbb{I}_{\theta=\theta'}$ applied to ϕ.

Next we show $\widehat{\phi}_A = \phi'_A$:

$$\begin{aligned}
\widehat{\phi}_A(a) &= \sum_{\theta\in\Theta} \widehat{\phi}(a,\theta) \\
&= \sum_{\theta\in\Theta}\sum_{a'\in A} \phi'(a|\eta_\phi(a'))\phi(a',\theta) \\
&= \sum_{a'\in A} \phi'(a|\eta_\phi(a'))\phi_A(a') \\
&= \sum_{\theta\in\Theta} \phi'(a|\theta)\phi'_\theta(\theta) = \phi'_A(a)
\end{aligned}$$

Finally, $E_{\widehat{\phi}}[\theta|a] = E_{\phi'}[\theta|a]$ for all a:

$$\begin{aligned}
\sum_{\theta\in\Theta} \theta\widehat{\phi}(\theta|a) &= \sum_{\theta\in\Theta}\theta\sum_{a'\in A} \phi'(a|\eta_\phi(a'))\phi(a',\theta)\frac{1}{\widehat{\phi}_A(a)} \\
&= \sum_{\theta\in\Theta}\theta\sum_{a'\in A} \phi'(\eta_\phi(a')|a)\frac{\phi(a',\theta)}{\phi'_\theta(\eta_\phi(a'))} \\
&= \sum_{a'\in A} \phi'(\eta_\phi(a')|a)\sum_{\theta\in\Theta}\theta\phi(\theta|a')
\end{aligned}$$

$$= \sum_{a' \in A} \phi'(\eta_\phi(a')|a)\eta_\phi(a') = \sum_{\theta \in \Theta} \theta \phi'(\theta|a)$$

□

To prove Proposition 4, we first note that when we pool together the conditional distributions of the state across the action profiles, the correlation between the state and action profile is reduced, and hence the f-information is also reduced:

Lemma 5. *For a $\phi \in \Delta(A \times \Theta)$, suppose $\beta = \phi_A$ places positive probability on a^1 and $a^2 \in A$. Let $\phi' \in \Delta(A \times \Theta)$ be such that $\phi'_A = \beta$ and*

$$\phi'(\theta \mid a) = \begin{cases} \phi(\theta \mid a) & a \notin \{a^1, a^2\}; \\ \phi(\theta \mid a^1)\frac{\beta(a^1)}{\beta(a^1)+\beta(a^2)} + \phi(\theta \mid a^2)\frac{\beta(a^2)}{\beta(a^1)+\beta(a^2)} & a \in \{a^1, a^2\}. \end{cases}$$

Then we have $D_f(\phi' \parallel \beta \otimes \mu) \leqslant D_f(\phi \parallel \beta \otimes \mu)$.

Proof. This follows from the DPI using the kernel:

$$b(a', \theta' \mid a, \theta) = \begin{cases} 1 & a = a' \notin \{a^1, a^2\} \text{ and } \theta = \theta' \\ \frac{\beta(a')}{\beta(a^1)+\beta(a^2)} & a', a \in \{a^1, a^2\} \text{ and } \theta = \theta' \\ 0 & \text{else} \end{cases}$$

□

Proof of Proposition 4. Suppose $\mu' = \alpha\tilde{\mu} + (1-\alpha)\widehat{\mu}$, where $\tilde{\mu}, \widehat{\mu} \in P_\mu$ and $\alpha \in [0,1]$ (for notational simplicity, we assume a convex combination of only two elements in P_μ). By the definition of P_μ, there exist $\tilde{\phi}, \widehat{\phi} \in \Delta(A \times \Theta)$ such that $\nu(\tilde{\phi}) = \tilde{\mu}$, $\nu(\widehat{\phi}) = \widehat{\mu}$, $D_f(\tilde{\phi} \parallel \tilde{\beta} \otimes \mu) \leqslant \epsilon$ and $D_f(\widehat{\phi} \parallel \widehat{\beta} \otimes \mu) \leqslant \epsilon$, where $\tilde{\beta}$ and $\widehat{\beta}$ are the marginal distributions of $\tilde{\phi}$ and $\widehat{\phi}$ over A, respectively. By Lemma 5, we can assume without loss that $\tilde{\phi}(a,\theta) = \tilde{\beta}(a)\tilde{\rho}(\theta \mid \tilde{\eta}(a))$ and $\widehat{\phi}(a,\theta) = \widehat{\beta}(a)\widehat{\rho}(\theta \mid \widehat{\eta}(a))$, where $\tilde{\eta}(a)$ and $\widehat{\eta}(a)$ are the interim state given a, and $\tilde{\rho}$ and $\widehat{\rho}$ are the unbiased noises conditional on the interim state.

Note that

$$\begin{aligned} D_f(\tilde{\phi} \parallel \tilde{\beta} \otimes \mu) &= \sum_{a \in A, \theta \in \Theta} f\left(\frac{\tilde{\phi}(a,\theta)}{\tilde{\beta}(a)\mu(\theta)}\right)\tilde{\beta}(a)\mu(\theta) \\ &= \sum_{a \in A, \theta \in \Theta} f\left(\frac{\tilde{\rho}(\theta \mid \tilde{\eta}(a))}{\mu(\theta)}\right)\tilde{\beta}(a)\mu(\theta), \end{aligned}$$

$$= \sum_{\theta\in\Theta,\theta'\in\Theta} f\left(\frac{\tilde{\rho}(\theta \mid \theta')}{\mu(\theta)}\right)\tilde{\mu}(\theta')\mu(\theta),$$

and analogously

$$D_f(\widehat{\phi} \parallel \widehat{\beta} \otimes \mu) = \sum_{\theta\in\Theta,\theta'\in\Theta} f\left(\frac{\widehat{\rho}(\theta \mid \theta')}{\mu(\theta)}\right)\widehat{\mu}(\theta')\mu(\theta),$$

Therefore, by convexity of f, we have

$$\begin{aligned}\epsilon &\geqslant \alpha D_f(\tilde{\phi} \parallel \tilde{\beta} \otimes \mu) + (1-\alpha) D_f(\widehat{\phi} \parallel \widehat{\beta} \otimes \mu) \\ &= \sum_{\theta\in\Theta}\sum_{\theta'\in\Theta} f\left(\frac{\tilde{\rho}(\theta \mid \theta')}{\mu(\theta)}\right)\alpha\tilde{\mu}(\theta')\mu(\theta) + \sum_{\theta\in\Theta}\sum_{\theta'\in\Theta} f\left(\frac{\widehat{\rho}(\theta \mid \theta')}{\mu(\theta)}\right)(1-\alpha)\widehat{\mu}(\theta')\mu(\theta) \\ &\geqslant \sum_{\theta\in\Theta}\sum_{\theta'\in\Theta} f\left(\frac{\rho'(\theta \mid \theta')}{\mu(\theta)}\right)\mu'(\theta')\mu(\theta),\end{aligned} \tag{20}$$

where

$$\rho'(\theta \mid \theta') = \tilde{\rho}(\theta \mid \theta')\frac{\alpha\tilde{\mu}(\theta')}{\mu'(\theta')} + \widehat{\rho}(\theta \mid \theta')\frac{(1-\alpha)\widehat{\mu}(\theta')}{\mu'(\theta')}.$$

Now, suppose $\phi' \in \mathrm{BCE}(\mu')$, and let $\mu'' = \nu(\phi')$. Effectively we need to show that $\mu'' \in P_\mu$. Without loss suppose there exists noise ρ'' such that $\phi'(a,\theta') = \beta'(a)\rho''(\theta' \mid \eta'(a))$ for all $a \in A$ and $\theta' \in \Theta$. Add noise ρ' to ϕ' to arrive at an outcome $\phi \in \mathrm{BCE}(\mu)$: $\phi(a,\theta) = \sum_{\theta'\in\Theta}\phi'(a,\theta')\rho'(\theta \mid \theta')$. Note that the marginal distributions of ϕ and ϕ' over A are the same: $\beta = \beta'$. We compute

$$\begin{aligned}D_f(\phi \parallel \beta \otimes \mu) &= \sum_{\theta\in\Theta}\sum_{\theta''\in\Theta} f\left(\frac{\sum_{\theta'\in\Theta}\rho''(\theta' \mid \theta'')\rho'(\theta \mid \theta')}{\mu(\theta)}\right)\mu''(\theta'')\mu(\theta) \\ &\leqslant \sum_{\theta\in\Theta}\sum_{\theta''\in\Theta}\sum_{\theta'\in\Theta} f\left(\frac{\rho'(\theta \mid \theta')}{\mu(\theta)}\right)\rho''(\theta' \mid \theta'')\mu''(\theta'')\mu(\theta) \\ &= \sum_{\theta\in\Theta}\sum_{\theta'\in\Theta} f\left(\frac{\rho'(\theta \mid \theta')}{\mu(\theta)}\right)\mu'(\theta')\mu(\theta) \\ &\leqslant \epsilon,\end{aligned}$$

where the last inequality follows from (20). Thus, we have $\mu'' \in P_\mu$, and $W(\phi') = W(\phi)$ is greater than or equal to the optimal value of problem (5).

Since ϕ' is an arbitrary element of $\mathrm{BCE}(\mu')$ and μ' is an arbitrary element of $\operatorname{conv} P_\mu$, we conclude that the optimal value of problem (7) is greater than or equal to that of problem (5).

Moreover, the optimal value of problem (7) is obviously less than or equal to that of the right-hand side of problem (5), which is also equal to the left-hand side by Theorem 5. This proves the proposition. □

B Relations on information structures

In this Appendix, we discuss the connection between our results and the literature on comparisons of information structures. The notions of individual garbling and coordinated individual garbling have been previously introduced in the literature, notably in Gossner (2000) and Lehrer, Rosenberg, and Shmaya (2013). We discuss each of these in turn.

Gossner (2000) asks the question: Given information structures I and I', when is it the case that $E_I(G) \subseteq E_{I'}(G)$ for all G? His main result shows that this is the case if and only if, in our terminology, I is a coordinated individual garbling of I'.[9] Thus, coordinated individual garblings represent a natural preorder on information structures.

Gossner's setup is different from ours in that he allows for compact and continuous games. As a result, his result does not imply ours, nor do our results imply his. We now state an analogue of Gossner's theorem for our finite setting:

Proposition 9. *I is a coordinated individual garbling of I' if and only if $E_I(G) \subseteq E_{I'}(G)$ for all G.*

Proof of Proposition 9. For the if direction, let G be the separation game and ϕ the equilibrium outcome, as in Lemma 1. By hypothesis, $\phi \in E_{I'}(G)$, so by Lemma 1, I is a coordinated individual garbling of I'.

We now prove the only if direction. Suppose $I = (S, \sigma)$ is a coordinated individual garbling of $I' = (S', \sigma')$, where $b' \in B(S, S')$ is the individual garbling that satisfies (2). Let $G = (A, u)$ and $\phi \in E_I(G)$, and b the strategies that induce ϕ. We claim that the strategies

$$\widehat{b}_i(a_i|s_i') = \sum_{s_i \in S_i} b_i(a_i|s_i) b_i'(s_i|s_i')$$

are an equilibrium of (I', G) that induces ϕ. To see that $\widehat{b}$ is an equilibrium, first note that

$$\begin{aligned}\widehat{b}_{-i}(a_{-i}|s_{-i}') &\equiv \prod_{j \neq i} \sum_{s_j \in S_j} b_j(a_j|s_j) b_j'(s_j|s_j') \\ &= \sum_{s_{-i} \in S_{-i}} b_{-i}(a_{-i}|s_{-i}) b_{-i}'(s_{-i}|s_{-i}'),\end{aligned}$$

[9]Gossner would say that I is "faithfully reproduced" from I'.

where analogously $b_{-i}(a_{-i}|s_{-i}) \equiv \prod_{j \neq i} b_j(a_j|s_j)$, and so on. Now, by equation (2), for every s_i and $s'_i \in S'_i$ such that $b'_i(s_i|s'_i) > 0$, there exists an $\alpha > 0$ such that for all (s_{-i}, θ),

$$\alpha\sigma(s_i, s_{-i}, \theta) = b'_i(s_i|s'_i) \sum_{s'_{-i}} \left(\prod_{j \neq i} b'_j(s_j|s'_j) \right) \sigma'(s'_i, s'_{-i}, \theta).$$

Hence, for any i, s'_i, and a'_i,

$$\begin{aligned} &\sum_{a_{-i}} \sum_{s'_{-i} \in S'_{-i}, \theta \in \Theta} \widehat{b}_{-i}(a_{-i}|s'_{-i})\sigma'(s'_i, s'_{-i}, \theta)u_i(a'_i, a_{-i}, \theta) \\ &= \sum_{a_{-i}} \sum_{s'_{-i} \in S'_{-i}, \theta \in \Theta} \sum_{s_{-i} \in S_{-i}} b_{-i}(a_{-i}|s_{-i})b'(s_{-i}|s_{-i})\sigma'(s'_i, s'_{-i}, \theta)u_i(a'_i, a_{-i}, \theta) \\ &= \alpha \sum_{a_{-i}} \sum_{s_{-i} \in S_{-i}, \theta \in \Theta} b_{-i}(a_{-i}|s_{-i})\sigma(s_i, s_{-i}, \theta)u_i(a'_i, a_{-i}, \theta), \end{aligned}$$

which is proportional to the interim expected utility of type s_i playing a'_i when others are using b_{-i}, in the information structure I. Thus, if a_i is a best response for some s_i with $b'(s_i|s'_i) > 0$, then it is a best response for s'_i as well. Since $\widehat{b}_i(\cdot|s'_i)$ is supported on such a_i's, and it follows that $\widehat{b}_i$ is a best response to $\widehat{b}_{-i}$. Finally, because b' is an individual garbling, we have that

$$\begin{aligned} \sum_{s' \in S'} \widehat{b}(a|s')\sigma'(s', \theta) &= \sum_{s \in S} b(a|s) \sum_{s' \in S'} b'(s|s')\sigma'(s', \theta) \\ &= \sum_{s \in S} b(a|s)\sigma(s, \theta) = \phi(a, \theta), \end{aligned}$$

so that $\widehat{b}$ and I' induce ϕ. We conclude that $\phi \in E_{I'}(G)$, as desired. □

The proof of the only if is morally the same as Gossner's: From the fact that I is a coordinated individual garbling of I, we know that the players can "simulate" I from I' in such a way that each player's garbled signal is sufficient for other players' garbled signals and the state. Thus, any equilibrium of I has an equivalent equilibrium of I', where the players first simulate I and then play the given strategies on I.

The if direction of the proof of Proposition 9, however, is substantively different. Gossner also constructs a game that plays an analogous role as our separation game. However, in his game, players report signals as well as beliefs about others' signals (and the state, in his extension to incomplete information). The payoff for the belief is given by a log scoring rule, with the payoff defined to be $-\infty$ if a player assigns zero probability to the signals that are reported by the others. Because Gossner's game is not compact, an extra step is needed to approximate the obedient outcome of his separation game via compact games. He establishes that an approximate analogue of the coordinated individual garbling property holds, and he finally takes limits to conclude that it holds exactly. In comparison, the construction and analysis of our separation game is elementary. The game is finite, and no scoring rules, infinite payoffs, or approximations are needed.

In a slightly different but related direction, Lehrer, Rosenberg, and Shmaya (2013) study equivalence relations on information structures that arise from having the same set of equilibrium outcomes for all games, according to various equilibrium concepts, and including Bayes Nash equilibrium. Their main result for Bayes Nash equilibrium is that two information structures have the same equilibrium outcomes for all games if and only if they are individual garblings of one another. While it is not a primary objective of our paper, our investigation has led us to the observation that the equivalence relations of Lehrer, Rosenberg, and Shmaya (2013) are also equivalent to several other notions, which we now state.

We say that $I' = (S', \sigma')$ is a *reduction* of $I = (S, \sigma)$ if there are mappings $f_i : S_i \to S'_i$ for each i such that (i) if $f_i(s_i) = f_i(\widehat{s}_i)$, then there exists an $\alpha \in \mathbb{R}$ such that for all $(s_{-i}, \theta) \in S_{-i} \times \Theta$, we have

$$\sigma(s_i, s_{-i}, \theta) = \alpha\sigma(\widehat{s}_i, s_{-i}, \theta),$$

and (ii) for any $(s', \theta) \in S' \times \Theta$, we have that

$$\sigma'(s', \theta) = \sum_{s \in f^{-1}(s')} \sigma(s, \theta).$$

In other words, I' is obtained from I by merging types that have the same interim beliefs. We say that I is *irreducible* if for every i, no two types have the same interim beliefs. We say that I is *reduction equivalent* to I' if there is an information structure I'' that is a reduction of both I and I'.

Proposition 10. *Given information structures I and I', the following statements are equivalent:*

(a) I and I' are individual garblings of each other.

(b) I and I' are coordinated individual garblings of each other.

(c) $E_I(G) = E_{I'}(G)$ for all G (equilibrium outcome equivalence).

(d) $F_I(A) = F_{I'}(A)$ for all A (outcome equivalence).

(e) I and I' are reduction equivalent.

Proof of Proposition 10. For the sake of completeness, we give a self-contained proof of the proposition, including those parts known from prior work.

(b) $\implies$ (a): Clearly, if I and I' are coordinated individual garblings of each other, then they are also individual garblings of each other.

(a) $\implies$ (d): Suppose $I = (S, \sigma)$ is an individual garbling of $I' = (S', \sigma')$, with garbling $b' \in B(S', S)$, and let $\phi \in F_I(A)$, induced by strategies $b' \in B(S, A)$. Define the strategies $\widehat{b} \in B(S', A)$ by

$$\widehat{b}_i(a_i|s'_i) = \sum_{s_i \in S_i} b_i(a_i|s_i) b'_i(s_i|s'_i).$$

Then the outcome ϕ' induced by $\widehat{b}$ and I' is

$$\begin{aligned}
\phi'(a, \theta) &= \sum_{s' \in S'} \widehat{b}(a|s')\sigma'(s', \theta) \\
&= \sum_{s' \in S'} \prod_{i=1,\ldots,N} \sum_{s_i \in S_i} b_i(a_i|s_i) b'_i(s_i|s'_i)\sigma'(s', \theta) \\
&= \sum_{s' \in S'} \sum_{s \in S} b(a|s) b'(s|s)\sigma'(s', \theta) \\
&= \sum_{s \in S} b(a|s) \sum_{s' \in S'} b'(s|s')\sigma'(s', \theta) \\
&= \sum_{s \in S} b(a|s)\sigma(s, \theta) \\
&= \phi(a, \theta)
\end{aligned}$$

so that $\phi \in F_{I'}(A)$. Since ϕ was arbitrary, we have $F_I(A) \subseteq F_{I'}(A)$. Reversing the roles of I and I' gives $F_{I'}(A) \subseteq F_I(A)$, so that I and I' are outcome equivalent, as desired.

(b) $\implies$ (c): This follows immediately from Proposition 9.

(c) $\implies$ (b): Let G be the separation game and ϕ the equilibrium outcome $\phi \in E_I(G)$ given in Lemma 1. By (c), $\phi \in E_{I'}(G)$, so by Lemma 1, I is a coordinated individual garbling of I'. Repeating this argument with I and I' reversed implies that I' is an individual garbling of I as well.

(c) $\implies$ (d): Fix an action space A and let $G = (A, u)$ where $u_i(a) = 0$ for all i and $a \in A$. From equilibrium outcome equivalence, we have that $E_I(G) = E_{I'}(G)$. But because players are indifferent between all actions, $E_I(G) = F_I(A)$ and $E_{I'}(G) = F_{I'}(A)$. We conclude that $F_I(A) = F_{I'}(A)$, i.e., I and I' are outcome equivalent.

(d) $\implies$ (e): Suppose that I and I' are outcome equivalent. Since I and I' are reduction equivalent to their respective reductions, it is without loss to assume that I and I' are irreducible. Now, clearly we have that $\sigma \in F_I(S)$ and $\sigma' \in F_{I'}(S')$. Outcome equivalence therefore implies that $\sigma \in F_{I'}(S)$ and $\sigma' \in F_I(S')$. Let b and b' be strategies such that (I, b) induce σ' and (I', b') induce σ. Define the Markov kernels $K_i : S_i \to \Delta(S_i)$ according to

$$K_i(\widehat{s}_i|s_i) = \sum_{s'_i \in S'_i} b'_i(\widehat{s}_i|s'_i) b_i(s'_i|s_i).$$

Also define the product kernel $K(\widehat{s}|s) = \prod_{i=1,\dots,N} K_i(\widehat{s}_i|s_i)$. It follows from the fact that (I, b) induce σ' and (I', b') induce σ that σ is an invariant measure for K, in the sense that for all $(\widehat{s}, \theta)$, we have

$$\sigma(\widehat{s}, \theta) = \sum_{s \in S} K(\widehat{s}|s)\sigma(s, \theta).$$

Now, let d_i such that the kernel $K_i^{d_i}$ is aperiodic, and let $d = \prod_{i=1,\dots,N} d_i$ (so that K_i^d and K^d are all aperiodic). Let P_i be the partition of S_i into communicating classes of K_i and let P be the partition of S into communicating classes of K. It is easy to see that $P = \prod_{i=1,\dots,N} P_i$. Note that because K_i^d is aperiodic, there is a unique invariant measure of K_i^d restricted to $p_i \in P_i$, which we denote by $\pi_i^{p_i}$. Similarly, if K^d is restricted to $p = \prod_{i=1,\dots,N} p_i \in P$, there is a unique invariant measure π^p, and since $\prod_{i=1,\dots,N} \pi_i^{p_i}$ is also an invariant of K^d on p, we must have $\pi^p = \prod_{i=1,\dots,N} \pi_i^{p_i}$. Hence, if we write

$$\sigma(p, \theta) = \sum_{s \in p} \sigma(s, \theta),$$

then for $s \in p \in P$, we have

$$\sigma(s, \theta) = \sigma(p, \theta)\pi^p(s).$$

Now, fix i, $s \in p \in P$, and $\theta \in \Theta$. Then

$$\begin{aligned}
\sigma_i(s_{-i}, \theta|s_i) &= \frac{\sigma(s_i, s_{-i}, \theta)}{\sum_{s'_{-i} \in S_{-i}, \theta' \in \Theta} \sigma(s_i, s'_{-i}, \theta')} \\
&= \frac{\pi^p(s_i, s_{-i})\sigma(p, \theta)}{\sum_{p'_{-i} \in P_{-i}, \theta' \in \Theta} \sigma(p_i, p'_{-i}, \theta') \sum_{s'_{-i} \in p'_{-i}} \pi^{p_i, p'_{-i}}(s_i, s'_{-i})} \\
&= \frac{\pi_i^{p_i}(s_i)\pi_{-i}^{p_{-i}}(s_{-i})\sigma(p, \theta)}{\sum_{p'_{-i} \in P_{-i}, \theta' \in \Theta} \sigma(p_i, p'_{-i}, \theta') \sum_{s'_{-i} \in p'_{-i}} \pi_i^{p_i}(s_i)\pi_{-i}^{p'_{-i}}(s'_{-i})}
\end{aligned}$$

$$= \frac{\pi_i^{p_i}(s_i)\pi_{-i}^{p_{-i}}(s_{-i})\sigma(p,\theta)}{\pi_i^{p_i}(s_i)\sum_{p'_{-i}\in P_{-i},\theta'\in\Theta}\sigma(p_i,p'_{-i},\theta')}$$
$$= \frac{\pi_{-i}^{p_{-i}}(s_{-i})\sigma(p,\theta)}{\sum_{p'_{-i}\in P_{-i},\theta'\in\Theta}\sigma(p_i,p'_{-i},\theta')}.$$

This expression depends on p_i but not on the particular $s_i \in p_i$. Hence, it must be that if $s_i, s'_i \in p_i$, then $\sigma_i(\cdot,\cdot|s_i) = \sigma_i(\cdot,\cdot|s'_i)$. From the hypothesis that I is irreducible, we conclude that $|p_i| = 1$, and thus $K^d(s|s) = 1$. This is possible only if b' is a pure strategy such that $b'(s|s') = 1$ for any s' such that $b(s'|s) > 0$. By a similar analysis, we conclude that b' is also pure. Hence, the function $f_i(s_i)$ defined according to $b_i(f_i(s_i)|s_i) = 1$ is a bijection from S_i to S'_i. This function satisfies (i) and (ii) in the definition of reduction, so that I' is a reduction of I, and vice versa.

(e) $\Longrightarrow$ (c): We will show this for the special case in which $S' \subseteq S$, and $f_i(s''_i) = s''_i$ for all $s''_i \notin \{s_i, s'_i\}$ and $f_j(s_j) = s_j$ for all $s_j \in S_j$. In other words, exactly two signals are merged for player i, and no signals are merged for players $j \neq i$. Without loss, we assume that s_i and s'_i arise with positive probability. Let $\sigma_i(s_i)$ denote the marginal probability of s_i, and let $\sigma_i(s_{-i},\theta|s_i)$ denote the conditional distribution of (s_{-i},θ) given s_i, and defined similarly for σ'. From the definition of reduction equivalence, we have that $\sigma'_j(\cdot,\cdot|f_j(s_j)) = \sigma_j(\cdot,\cdot|s_j)$ and $\sigma'_j(s'_j) = \sum_{s_j\in f_j^{-1}(s'_j)}\sigma_j(s_j)$ for all j and $s_j \in S_j$.

Now, let b and b' be strategies in (I,G) and (I',G) respectively, such that (i) $b'_j(s_j) = b_j(s_j)$ for all $j \neq i$ and s_j, (ii) $b'_i(s''_i) = b_i(s''_i)$ for all $s''_i \notin \{s_i, s'_i\}$, and (iii)

$$b'_i(\bar{s}_i) = \frac{1}{\sigma_i(s_i)+\sigma_i(s'_i)}\left(\sigma_i(s_i)b_i(s_i) + \sigma_i(s'_i)b_i(s'_i)\right). \tag{21}$$

where $\bar{s}_i = f_i(s_i) = f_i(s'_i)$. Then (I',b') induce the outcome

$$\sum_{s'\in S}\sigma'(s',\theta)b'(a|s')$$
$$= \sum_{s_{-i}\in S_{-i}}\left[\sum_{s''_i\notin\{s_i,s'_i\}}\sigma'(s''_i,s_{-i},\theta)b'(a|s''_i,s_{-i}) + \sigma'_i(\widehat{s}_i)\sigma'(s_{-i},\theta|\widehat{s}_i)b'(a|\widehat{s}_i,s_{-i})\right]$$
$$= \sum_{s_{-i}\in S_{-i}}\left[\sum_{s''_i\notin\{s_i,s'_i\}}\sigma(s''_i,s_{-i},\theta)b(a|s''_i,s_{-i}) + \sigma'_i(\widehat{s}_i)b'_i(a_i|\widehat{s}_i)\sigma'(s_{-i},\theta|\widehat{s}_i)\prod_{j\neq i}b'_{-i}(a_{-i}|s_{-i})\right]$$
$$= \sum_{s_{-i}\in S_{-i}}\left[\sum_{s''_i\notin\{s_i,s'_i\}}\sigma(s''_i,s_{-i},\theta)b(a|s''_i,s_{-i}) + \left(\sum_{s''_i\in\{s_i,s'_i\}}\sigma_i(s''_i)b'_i(a_i|s''_i)\right)\sigma(s_{-i},\theta|s''_i)\prod_{j\neq i}b_{-i}(a_{-i}|s_{-i})\right]$$

$$= \sum_{s_{-i}\in S_{-i}} \left[\sum_{s_i''\notin\{s_i,s_i'\}} \sigma(s_i'', s_{-i}, \theta) b(a|s_i'', s_{-i}) + \left(\sum_{s_i''\in\{s_i,s_i'\}} \sigma_i(s_i'') b_i(a_i|s_i'') \right) \sigma(s_{-i}, \theta|s_i'') \prod_{j\neq i} b_{-i}(a_{-i}|s_{-i}) \right]$$
$$= \sum_{s\in S} \sigma(s,\theta) b(a|s) = \phi(a,\theta).$$

Hence $U_i(b'; I', G) = U_i(b; I, G)$.

Now, we claim that if b and b' satisfy (i)–(iii) above, then b is an equilibrium if and only if b' is an equilibrium. We have already established that $U_j(b; I, G) = U_j(b'; I', G)$ for all j. If b' is not an equilibrium, then there exists j and $\widehat{b}_j'$ such that $U_j(\widehat{b}_j', b_{-j}'; I', G) > U_j(b'; I', G)$. Now define $\widehat{b}_j(s_j) = \widehat{b}_j'(f_j(s_j))$. Then $(\widehat{b}_j', b_{-j}')$ and $(\widehat{b}_j, b_{-j})$ satisfy (i)–(iii), so that $U_j(\widehat{b}_j, b_{-j}; I, G) = U_j(\widehat{b}_j', b_{-j}'; I', G)$, so that $\widehat{b}_j$ is a profitable deviation from (I, b). Alternatively, if $\widehat{b}_j$ is a profitable deviation from (I, b), then we can define $\widehat{b}_j'$ according to (21), so that (i)–(iii) are again satisfied for $(\widehat{b}_j', b_{-j}')$ and $(\widehat{b}_j, b_{-j})$, and we similarly conclude that $\widehat{b}_j'$ is a profitable deviation from (I', b'). This completes the proof that I and I' are equilibrium outcome equivalent when I' is a reduction of I obtained by merging two signals.

Iterative application of this step (and relabeling of signals) then establishes that I is equilibrium outcome equivalent to I' if I' is a reduction of I. Now, if I and I' are reduction equivalent, then there exists an information structure I'' that is a reduction of I and I'. We therefore have $E_I(G) = E_{I''}(G) = E_{I'}(G)$.

We conclude that $(b) \iff (c) \implies (d) \implies (e) \implies (c)$. So that (b)–(e) are all equivalent. Finally, $(a) \implies (d)$ and $(b) \implies (a)$, so that (b)–(e) are all equivalent to (a) as well. □

The proof of Proposition 10 shows that if I' is a reduction of I and I'' is a reduction of I', then I'' is in fact a reduction of I. This implies that any for information structure I, there is a unique I' (up to a relabeling of signals) that is an irreducible reduction of I. Moreover, I' can be obtained "at one step", by merging signals in I that have the same interim belief. In this sense, there is a simple finite procedure for determining if I and I' are equivalent, that eliminates the existential and universal quantifiers over individual garblings and games, respectively.