Implementability of Correlated and Communication Equilibrium Outcomes in Incomplete Information Games

Igal Milchtaich*

October 2010

Abstract
In a correlated equilibrium, the players' choice of actions is directed by random, correlated messages that they receive from an outside source, or mechanism. This allows for more equilibrium outcomes than without such messages (pure-strategy equilibrium) or with statistically independent ones (mixed-strategy equilibrium). In an incomplete information game, the messages may also convey information about the types of the other players, either because they reflect extraneous events that affect the types (correlated equilibrium) or because the players themselves report their types to the mechanism (communication equilibrium). Thus, mechanisms can be classified by the connections between the messages that the players receive and their own and the other players' types, the dependence or independence of the messages, and whether or not randomness is involved. These properties may affect the achievable equilibrium outcomes, i.e., the payoffs and joint distributions of type and action profiles. Whereas for complete information games there are only three classes of equilibrium outcomes, with incomplete information the number is 14–15 for correlated equilibria and 15–17 for communication equilibria. Each class is characterized by the properties of the mechanisms that implement its members. The majority of these classes have not been described before. *JEL Classification:* C72.

Keywords: Correlated equilibrium; Communication equilibrium; Incomplete information; Bayesian games; Mechanism; Correlation device; Implementation

1 Introduction

A pure-strategy Nash equilibrium in a complete information strategic game represents a possible outcome for rational players who do not randomize over actions. Adding the possibility to randomize extends the set of equilibrium outcomes by facilitating mixed-strategy equilibria. For a correlated equilibrium, independent randomization devices are not sufficient – an external correlation device is required. Thus, the set of feasible equilibrium outcomes expands as increasingly richer mechanisms are allowed. A similar relation between equilibrium outcomes and mechanisms holds for incomplete information games. However, the relation in this case is more complex than for complete information games. This is because the set of equilibrium outcomes implementable by a mechanism depends on the extent to which its output reflects the players' types. Implementability may depend, for

* Department of Economics, Bar-Ilan University, Ramat Gan 52900, Israel
milchtaich@ biu.ac.il http://faculty.biu.ac.il/~milchti

example, on whether the messages that the mechanism sends to the players provide them with information about the other players' types, and on whether messages depend on the receiver's type. The former affects the mechanism's ability to implement type-dependent coordinated actions, and the latter affects its ability to transmit information selectively, that is, to certain types of players only.

A mechanism may facilitate type-dependent coordinated actions if it is affected by factors that also affect the players' types. For example, whether the economy is booming of slumping may affect both the types of firms that enter an auction and the various macroeconomic indicators (e.g., the CPI) that these firms factor in when deciding on their respective bids. By contrast, for selectivity, "knowledge" of the players' types is not necessary: type-dependent perceptual abilities may suffice. For example, different recommendations may be issued to unilingual English and French readers simply by handing out a bilingual sheet with English and French texts that do not match.

A simple, straightforward way to implement type-dependent coordinated actions is to ask the players to report their types. However, for this to work, the actions have to be such that truthful type reports are incentive compatible. This requirement defines communication equilibrium and distinguishes it from correlated equilibrium, which only requires the actions to be incentive compatible and where communication is one-way – messages from the mechanism to the players.

The main objective of this paper is to chart the connections between properties of mechanisms and the kinds of correlated and communication equilibrium outcomes implementable by them in incomplete information games. These connections form a rich and intricate structure, and they are not always obvious or perfectly intuitive. The subject matter is quite different – both in substance and in the relevant techniques – from issues studied in the context of complete information games. It has some similarity to the problem of implementability of social choice functions studied in mechanism-design theory, which is reflected by the similar terminology. However, implementability of correlated and communication equilibrium outcomes is not a special case of implementability of social choice functions (see also Kar et al., 2010).

A second, auxiliary objective of the paper is to present a single framework that accommodates the majority of the previously described varieties of correlated strategies, correlated equilibria and communication equilibria in incomplete information games, in that each variety corresponds to a particular set of structural limitations on the allowed mechanisms. The power of these limitations stems from their effect on the mechanisms' ability to orchestrate certain joint actions, make the actions incentive incompatible, or elicit truthful type reports.

The paper's plan of attack is to separate the implementability problem into three interrelated problems. The first problem is the implementability of correlated strategy distributions. Here, only the joint distribution of the players' types and actions matters, and payoffs are irrelevant. The second problem, which does take payoffs and incentive compatibility into account, is the implementability of correlated and communication equilibrium distributions. The third and ultimate problem is the implementability of payoff vectors. The payoffs are uniquely determined by the joint distribution of types and actions, but not conversely. The advantages of this three-part approach in comparison with directly addressing the third problem are that it makes certain issues significantly more manageable and provides insights about the roots of non-implementability where it occurs.

The presentation of the results is divided into two parts. Following the layout of the basic framework in Section 2, Section 3 gives an overview of the results, mainly in the form of Hesse diagrams that present the different classes of correlated strategy outcomes, correlated equilibrium outcomes and communication equilibrium outcomes and the connections between the various classes. Section 3 also includes several examples that illustrate the results and the various issues involved. The subsequent four sections give the details and the proofs, and Section 8 summarizes.

1.1 Related literature

Aumann (1987) demonstrated that correlated equilibrium can be viewed as an expression of Bayesian rationality. A rational player's choice of action reflects his knowledge of the state of the world. The state includes a specification of the knowledge of the other players, which determines their actions. Bayesian rationality means that each player's action is a best reply to what he knows about the others' actions. Aumann's paper only concerned complete information games; types of players and type-dependent payoffs are not part of the setting. However, since the state-space formulation is a standard model for Bayesian games, the paper pointed to the logical next step, which was to merge the two settings by allowing the states of the world to determine the players' types in addition to any information that they posses which may be used for choosing an action. Crucially, this additional information is not specified by the game – it is part of the solution concept.

Two models of this kind were proposed by Cotter (1991, 1994), which differ from one another in the restrictions they put on the players' information. In a strategy correlated equilibrium (Cotter, 1991), the additional information takes the form of random messages that the players receive from an outside correlation mechanism, which is ignorant of their types. A correlated strategy with such a mechanism is a rule that maps the message each player receives to a strategy for that player, which is a prescription of a pure or randomized action for each of the player's types. The equilibrium condition is that acting accordingly is incentive compatible in that no player can increase his expected payoff by associating different strategies with the messages he receives. A type correlated equilibrium (Cotter, 1994; see also Samuelson and Zhang, 1989) can be described as a strategy correlated equilibrium in a version of the game in which each type of each player is an independent agent. This means that the mechanism sends to each player not a complete strategy but only the action it prescribes to the player's actual type. The message may thus depend on the player's type, unlike in a strategy correlated equilibrium, but it is still unaffected by the other players' types. Consequently, the player's action is conditionally independent of the other players' types, given the player's own type. Cotter stated that this so-called conditional independence property of the joint distribution of types and actions is characteristic of type correlated equilibria, in that any distribution with this property can be implemented by a mechanism as above. However, it was later shown that this assertion is incorrect (Milchtaich, 2004, Example 6).

A different extension of correlated equilibrium to games with incomplete information is communication, or mediated, equilibrium (Myerson, 1994). This solution concept differs from those considered above in that communication is two-way. The players first send private messages to, and then receive such messages from, a particular mechanism, which thus serves as a mediator as well as a correlation device. According to the revelation principle (see Myerson, 1994), without loss of generality the messages sent by the players to the mediator may be assumed type reports. The message that each player gets from the mediator indicates a particular action for that player. This mechanism in required to be incentive compatible in that it is in each player's best interest to report his type honestly and take the indicated action if all the others do the same.

The most comprehensive account to date of correlated and communication equilibria in games with incomplete information is Forges' (1993) paper, which compared strategy correlated equilibrium, type (or agent normal form) correlated equilibrium, communication equilibrium, and 'Bayesian solution'. (A fifth solution concept considered in the paper concerns hierarchies of beliefs.) Bayesian solution is a very general solution concept, which includes strategy and type correlated equilibria as special cases. It extends an incomplete information game by introducing a state space in which several states may correspond to a single type profile. This allows players to have partial or complete information about the other players' types as well as about outside events. As in Aumann's (1987) model, the information structure is complemented by a mapping from states to action profiles that is required to satisfy the obvious incentive compatibility condition. A Bayesian solution may be implemented by an omniscient mediator, who knows the players' types. In this, it differs from a communication equilibrium, in which the mediator totally relies on the players' type reports.

The messages that the players receive from the mediator are part of the solution concept, and are distinct from any signals that are part of the game itself and *define* the players' types. The potential impact of the former (the mediator's messages) depends on the degree of dependence among the latter (the players' types). For example, with perfectly correlated types, the players already know each other's type when they receive the mediator's messages, which can therefore only help them coordinate their actions. Conversely, if the types are independent, the mediator's messages may also inform players about the other players' types. However, this may be so only if the solution concept allows the messages to depend on the others' types. Therefore, depending on the solution concept, garbling, or randomly perturbing, in a particular way the signals that the players receive as part of the game may or may not change the set of equilibrium outcomes.[1] Lehrer et al. (2006) identified the kinds of garbling that do not affect the equilibrium outcomes for three kinds of correlated equilibrium in two-player Bayesian games: mixed (Nash) equilibrium, type correlated equilibrium, and a special kind of Bayesian solution (called belief invariant Bayesian solution by Forges, 2006), which satisfies a condition similar to the conditional independence property. They showed, for example, that garbing has no effect on mixed

[1] This assumes that the players' types do not have a *direct* effect on payoffs but only reflect certain effective "hidden variables", so that garbling the signals also has no direct effect on payoffs.

equilibria, regardless of the payoff functions, if and only if it is performed independently for each player, that is, without taking into account the other player's signal.

Identification of information types (Milchtaich, 2004) is a special kind of garbling. It removes distinctions between player types that are interchangeable in terms of their effect on the player's own payoff and those of the others and differ, say, only in what they knows about the other players' types. Identification of information types may transform one kind of equilibrium into another. For example, a pure (-strategy) equilibrium (with different actions for different information types) may become a mixed equilibrium (with several possible actions for the single type that results from the identification). Thus, the collection of pure equilibria, for example, is not closed under identification of information types. The same is true for more general solution concepts. For example, this is so for type correlated equilibrium, since when information types are identified, the conditional independence property may cease to hold (Milchtaich, 2004, Examples 7). In fact, the narrowest extension of pure equilibrium that is closed under identification of information types is the notion of correlated equilibrium used in the present paper (Milchtaich, 2004, Propositions 4 and 5), which is similar to Forges' (1993, 2006) Bayesian solution, or global equilibrium in the terminology of Lehrer et al. (2006). Thus, in this respect at least, this solution concept is not excessively broad.

2 Preliminaries

2.1 Bayesian games

An n-player (finite) *pre-Bayesian game* is a function $u = (u_1, u_2, \dots, u_n): T \times A \to \mathfrak{R}^n$, where $T = T_1 \times T_2 \times \cdots \times T_n$ and $A = A_1 \times A_2 \times \cdots \times A_n$ are each the Cartesian product of n finite sets and $\mathfrak{R}$ is the real line. The interpretation is that $N = \{1,2, \dots, n\}$ is the set of *players*; for each player i, the sets T_i and A_i and the function u_i are respectively the *type space*, *action space* and *payoff function* of player i; and for each *type profile* $t = (t_1, t_2, \dots, t_n) \in T$ and *action profile* $a = (a_1, a_2, \dots, a_n) \in A$, $u(t, a)$ is the *payoff vector*. A somewhat more general possibility is that the type of one of the players represents the *state of nature*, i.e., variables that affect the (real) players' payoffs but are unknown to them. That player, Nature, does not take any action, so that his action space must be a singleton.

A *pure strategy* for a player i is a function from the player's type space to his action space, i.e., an element of $A_i^{T_i}$. Using some fixed indexing of the (finite) type space, $T_i = \{t_i^1, t_i^2, \dots\}$, any pure-strategy can be written as $(a_i^1, a_i^2, \dots)$, where, for each j, a_i^j is the action prescribed to the jth type of player i. A *pure-strategy profile* is an assignment of a pure strategy to each player, i.e., an element of $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$, which can be written as $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$.

A pre-Bayesian game becomes a (finite) *Bayesian game* when it is coupled with a specified probability measure on T, which assigns a probability to each type profile t. This measure η_T gives the distribution of type profiles in the game. Its support, $\mathrm{supp}(\eta_T)$, which is the collection of all type profiles t that have positive probability, may be a proper subset of T. However, it is assumed (essentially without loss of generality) that every type t_i of every player i is supported, in the sense that $(t_i, t_{-i}) \in \mathrm{supp}(\eta_T)$ for some *partial type profile* $t_{-i} = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$. The measure η_T can always be expressed as the distribution of a *random type profile*, which is a random variable[2] $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \dots, \boldsymbol{t}_n)$ with values in T such that

$$\eta_T(\{t\}) = \Pr(\boldsymbol{t} = t), \qquad t \in T.$$

For example, a random type profile can be constructed by restricting the measure η_T to its support and defining $\boldsymbol{t}$ as the identity map on $\mathrm{supp}(\eta_T)$. It does not normally matter which random type profile is used to express a particular distribution of type profiles, and in this paper the symbol $\boldsymbol{t}$ and the definite article are used for referring to any random type profile.

For each player i, the conditional distribution of $\boldsymbol{t}$, given the player's type $\boldsymbol{t}_i$, is interpreted as the *posterior beliefs* of player i about the players' types. The interpretation entails that each player i knows his own type t_i but does not necessarily know the types of the other players.

2.2 Mechanisms

A mechanism for a Bayesian game is an extraneous source of messages[3], which the players receive before they have to choose their actions. The messages that each player i may receive are elements of some finite set M_i, the player's (received) *message space*. For each type profile t, the profile of messages is given by a random variable $\boldsymbol{m}(t) = (\boldsymbol{m}_1(t), \boldsymbol{m}_2(t), \dots, \boldsymbol{m}_n(t))$ with values in the product set $M = M_1 \times M_2 \times \dots \times M_n$. Thus, a mechanism is specified by a random variable (specifically, a vector whose entries are indexed by the type profiles and are themselves random n-tuples) $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ that is independent of the random type profile $\boldsymbol{t}$. The independence assumption means that, given the type profile, any residual randomness in the messages is extraneous to the game (see Section 2.2.1 for further discussion of this point). However, it does not mean that the messages themselves are unrelated to the type profile. To take an extreme example, if the mechanism is an outside observer who is capable of finding out the players' types, the messages may fully convey that information:

$$\boldsymbol{m}_i(t) = t, \qquad i \in N, t \in T. \tag{1}$$

In this example, the mechanism is purely a source of information about the other players' types. Other mechanisms may serve primarily or exclusively as randomization devices, and convey little or no information about the types. A finer, more exact classification of mechanisms is facilitated by the following list of properties. Each property is expressed by a condition that the messages sent by the mechanism are required to satisfy for every player i

[2] A *random variable*, in this work, is any function from a finite probability space where each point has positive probability to an arbitrary set. Random variables are denoted by boldface letters and their arguments are always suppressed. Since the range of a function as above is finite, it would be possible in principle to restrict attention to real-, or even integer-, valued random variables. However, in practice such a restriction would be inconvenient.

[3] The term 'messages' is used here rather than 'signals' to emphasize the assumption that the sending mechanism is part of a solution concept rather than the game. 'Signals' in an incomplete information game are often synonymous with the players' types, which *are* part of the game.

and pair of type profiles $t = (t_i, t_{-i})$ and $t' = (t'_i, t'_{-i})$. The symbol $\overset{d}{=}$ denotes equality *in distribution*.

(S) Player i's type does not affect the message he receives:

$$\boldsymbol{m}_i(t) = \boldsymbol{m}_i(t'_i, t_{-i}).$$

($\tilde{S}$) Player i's type does not affect the distribution of the message he receives:

$$\boldsymbol{m}_i(t) \overset{d}{=} \boldsymbol{m}_i(t'_i, t_{-i}).$$

(O) The other players' types do not affect the message player i receives:

$$\boldsymbol{m}_i(t) = \boldsymbol{m}_i(t_i, t'_{-i}).$$

($\tilde{O}$) The other players' types do not affect the distribution of the message:

$$\boldsymbol{m}_i(t) \overset{d}{=} \boldsymbol{m}_i(t_i, t'_{-i}). \quad (2)$$

(D) The messages are non-random:

$$\boldsymbol{m}(t) \text{ has a degenerate distribution.}^{4} \quad (3)$$

(I) The messages that different players receive are (statistically) independent:

$$\boldsymbol{m}_1(t), \boldsymbol{m}_2(t), \ldots, \boldsymbol{m}_n(t) \text{ are independent.} \quad (4)$$

Conditions $\tilde{S}$ and $\tilde{O}$ require equality between two distributions, for specified t and t'. Since the players' types are actually random, this translates into equality between *conditional* distributions. Specifically, $\tilde{S}$ entails that the message that player i receives is conditionally independent of his type $\boldsymbol{t}_i$, given the other players' types $\boldsymbol{t}_{-i}$. In property $\tilde{O}$, $\boldsymbol{t}_i$ and $\boldsymbol{t}_{-i}$ are interchanged.

Equality between distributions is a weaker requirement than equality between the random variables themselves, as required by properties S, O and D.[5] The latter means that the random variables are equal with probability 1 (equivalently, pointwise equal). The distinction between equality in distribution and equality with probability 1 does not seem to have received sufficient attention in the existing literature on games with incomplete information. This paper shows that it has significant implications for correlated strategies and equilibria.

[4] A distribution is *degenerate* if all the probabilities are 0 or 1.

[5] For example, if a device satisfies $\tilde{O}$ and the types of all but two players are changed, the distribution of the message that each of these two players receives does not change. However, the *joint* distribution of the two messages may change, for example, uncorrelated messages may become correlated. By contrast, O would imply that the joint distribution does not change when the other players' types change. A subtler – but highly consequential – difference between O and $\tilde{O}$ also applies to two-player games. See footnote 16.

2.2.1 Independence lemma and the canonical mechanism

The definition of mechanism in effect assumes that the randomness or uncertainty regarding the messages that the players receive has two independent potential sources: the inherent randomness of the type profile, which the messages may reflect, and residual randomness, which persists also with a specified type profile. The following lemma shows that this independence assumption involves no loss of generality. Any joint distribution of types and messages can be produced by "mixing" the random type profile $\boldsymbol{t}$ with a random variable of the form $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ that is independent of $\boldsymbol{t}$. The outcome of the mixture is the random profile of messages $\boldsymbol{m}(\boldsymbol{t})$, the value of which is determined by first determining the realization of $\boldsymbol{t}$, which is a type profile t, and then determining the realization of $\boldsymbol{m}(t)$.[6]

Lemma 1. Let $M = M_1 \times M_2 \times \cdots \times M_n$ be any finite product set and η any probability measure on $T \times M$ whose marginal on T is equal to η_T, the distribution of type profiles. There exists a random variable $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ that is independent of the random type profile $\boldsymbol{t}$ such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{m}(\boldsymbol{t})$ is equal to η.

Proof. For any η as above, it is possible to construct for each type profile t a random variable $\boldsymbol{m}(t) = (\boldsymbol{m}_1(t), \boldsymbol{m}_2(t), \dots, \boldsymbol{m}_n(t))$ with values in M such that, first, for $t \in \operatorname{supp}(\eta_T)$ the distribution of $\boldsymbol{m}(t)$ is the probability measure on M that assigns to each element m the (conditional) probability

$$\frac{\eta(\{(t, m)\})}{\eta_T(\{t\})};$$

second,

for $t \notin \operatorname{supp}(\eta_T)$, $\boldsymbol{m}(t)$ satisfies (4), and in addition,
for every $i \in N$, (2) holds for some $t' \in T$ with $(t_i, t'_{-i}) \in \operatorname{supp}(\eta_T)$; (5)

and third,

$$\{\boldsymbol{m}(t)\}_{t \in T} \text{ are independent} \tag{6}$$

and are collectively independent of $\boldsymbol{t}$. It follows from the first and third requirements that

$$\Pr(\boldsymbol{t} = t, \boldsymbol{m}(\boldsymbol{t}) = m) = \eta_T(\{t\}) \Pr(\boldsymbol{m}(t) = m) = \eta(\{(t, m)\}), \qquad t \in T, m \in M.$$

Thus, the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{m}(\boldsymbol{t})$ is equal to η. ∎

The proof of Lemma 1 details the construction of a specific mechanism $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ for any given measure η as in the lemma. This *canonical mechanism*[7] is such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{m}(\boldsymbol{t})$ is equal to η, and in addition, (5) and (6) hold. The latter two

[6] An equivalent definition of the random profile of messages is $\boldsymbol{m}(\boldsymbol{t}) = \rho(\boldsymbol{t}, \boldsymbol{m})$, where the auxiliary function $\rho: T \times M^T \to M$ is defined by $\rho(t', (m(t))_{t \in T}) = m(t')$, i.e., it is the projection of the second argument on the coordinate specified by the first one. This function is thus the "vessel" in which the two independent ingredients $\boldsymbol{t}$ and $\boldsymbol{m}$ are mixed to produce the actual messages to the players.

[7] The proof of the lemma allows some latitude in the construction, which means that the canonical device may actually have more than one version. However, differences between versions are inconsequential.

special properties may seem purely technical. Property (5) concerns type profiles t that cannot actually occur (that is, $\Pr(\boldsymbol{t} = t) = 0$), and (6) concerns relations between different type profiles, which by definition cannot coexist (since only one type profile is realized). However, property (6) is quite potent in that it essentially prevents the mechanism from satisfying S or O, unless it also satisfies D. This is because two random variables that are equal *and* independent necessarily have degenerate distributions. Mainly because of this limitation, it is not possible to restrict attention to canonical mechanisms. They are, however, quite useful technical constructs.

2.3 Correlated strategies

With a specified mechanism, a *correlated strategy* $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ in an n-player Bayesian game specifies the action a_i that each player i takes as a function $\sigma_i: T_i \times M_i \to A_i$ of the player's type t_i and the message he receives m_i. Thus, $a_i = \sigma_i(t_i, m_i)$.[8] As indicated, the messages are part of the specification of the correlated strategy rather than the game. Their (potential) randomness and that of the types means that the actions are also random. The *random action profile* corresponding to a correlated strategy σ with a mechanism $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ is the random variable $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ defined by

$$\boldsymbol{a}_i = \sigma_i(\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})), \qquad i \in N. \tag{7}$$

Correlated strategies include several other kinds of strategies as special cases. If the message that each player receives is unaffected by the player's own type or the other players' types and is also non-random, i.e., if the mechanism satisfies S, O and D (which effectively means that the player does not receive any messages at all), then, for some fixed $m_1, m_2, \dots, m_n$,

$$\boldsymbol{a}_i = \sigma_i(\boldsymbol{t}_i, m_i), \qquad i \in N.$$

In this case, for each player i, σ_i associates a (deterministic) action a_i with each type t_i, which means that the correlated strategy is effectively a pure-strategy profile, and may be referred to as such. A more general case is that of a *mixed-strategy profile*, which is defined as a correlated strategy with a mechanism that satisfies S, O and I. These properties of the mechanism mean that the messages are independent and equal to $\boldsymbol{m}_1(t'), \boldsymbol{m}_2(t'), \dots, \boldsymbol{m}_n(t')$, where t' is any fixed type profile, and the random action profile satisfies

$$\boldsymbol{a} = (\sigma_1(\boldsymbol{t}_1, \boldsymbol{m}_1(t')), \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(t')), \dots, \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(t'))). \tag{8}$$

This is effectively the same as (and it can be implemented by) private randomization over

[8] By assumption, the specification of the actions is deterministic: randomized actions are not allowed. This assumption involves no loss of generality, and in particular, it does not exclude mixed strategies. It just means that even private randomizations are viewed as parts of a single large device, which may or may not be a physical entity. As an example of the former possibility, randomization may be relegated to the device from which the players receive messages. In this case, a player-specific random number is appended to each message, such that these n random numbers are independent. Clearly, a device modified in this way does not generally satisfy D, but the modification has no effect on properties $S, \tilde{S}, O, \tilde{O}$ or I.

pure strategies independently for each player.[9] An even more general case is that of a *random (pure-strategy) profile*, which is defined as a correlated strategy with a mechanism that satisfies S and O. The actions can still be presented as in (8), but since property I is not assumed, the randomization cannot generally be emulated by independent private randomizations.

For a specified correlated strategy σ with a mechanism $\boldsymbol{m}$, the joint distribution of the random type profile $\boldsymbol{t}$ and the actions $\boldsymbol{a}$ (the latter given by (7)) is called the *correlated strategy distribution* (CSD). It is of course possible for several pairs of strategies and mechanisms to have the same CSD. A CSD is said to be *implementable* by a particular mechanism (which *implements* the distribution) if there is *some* correlated strategy with that mechanism that gives the distribution. A CSD is a *pure-strategy distribution*, *mixed-strategy distribution* or *random-profile distribution* if it is implementable by some mechanism $\boldsymbol{m}$ with properties S, O and D, properties S, O and I, or properties S and O, respectively.

Every CSD η is a probability measure on $T \times A$ whose marginal on T is η_T.[10] Hence, it has a canonical mechanism $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ (see Section 2.2.1), in which the message space M_i of each player i is his action space A_i. The interpretation is that the canonical mechanism explicitly tells each player which action to take. The *canonical strategy* σ with this mechanism simply instructs the players to obey, that is, it is defined by

$$\sigma_i(t_i, m_i) = m_i, \qquad i \in N. \tag{9}$$

The corresponding random action profile is the *canonical random action profile*, which is given by

$$(\sigma_1(\boldsymbol{t}_1, \boldsymbol{m}_1(\boldsymbol{t})), \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(\boldsymbol{t})), \dots, \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(\boldsymbol{t}))) = \boldsymbol{m}(\boldsymbol{t}).$$

This equality proves that the canonical mechanism $\boldsymbol{m}$ of the CSD η implements η. Obviously, a similar equality holds for the canonical mechanism and strategy of *every* probability measure on $T \times A$ with the marginal η_T. Therefore, every such measure is a CSD. This establishes the following.

Lemma 2. In a Bayesian game, a necessary and sufficient condition for a probability measure on $T \times A$ to be a correlated strategy distribution is that the marginal on T is equal to the distribution of type profiles.

2.4 Correlated equilibria

The players' incentives in a Bayesian game are embodied by their payoff functions, $u_1, u_2, \dots, u_n$. For a correlated strategy σ with a mechanism $\boldsymbol{m}$, the payoff of each player i is

[9] A mixed-strategy profile may also be viewed as a *behavior strategy* for each player, that is, a randomized action for each of the player's types.

[10] Since the distribution of type profiles is given as part of the specification of the game, a CSD may also be viewed as an assignment of a probability measure on A to every type profile t, namely, the distribution of the players' action when they have types t (which is arbitrarily if t lies outside the support of η_T).

the random variable $u_i(\boldsymbol{t}, \boldsymbol{a})$, where $\boldsymbol{t}$ is the random type profile and $\boldsymbol{a}$ is the random action profile given by (7). The correlated strategy is incentive compatible if none of the players i can increase his expected payoff by a unilateral deviation, that is, by changing the function determining his action from σ_i to some other function $\sigma_i': T_i \times M_i \to A_i$. Equivalently, incentive compatibility means that the action that the correlated strategy specifies for each player always maximizes the conditional expectation[11] of the player's payoff, given his type and the message he receives. The latter condition is used in the following definition.

Definition 1. In a Bayesian game, a correlated strategy σ with a mechanism $\boldsymbol{m}$ is a *correlated equilibrium* if the corresponding random action profile $\boldsymbol{a}$ is such that, for every player i and action a_i' for that player,

$$E\big(u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i}))\big|\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})\big) \geq 0. \tag{10}$$

If a correlated strategy is a correlated equilibrium, then the correlated strategy distribution is said to be a *correlated equilibrium distribution* (CED). A CED is *implementable* by a particular mechanism if there is some correlated equilibrium with that mechanism that gives the distribution. A CED is a *pure-* or *mixed-equilibrium distribution* if it is implementable by some mechanism with properties S, O and D, or properties S, O and I, respectively. Two additional kinds of CEDs are defined in Section 3.2.3.

An equivalent definition of CED, which does not explicitly refer to an implementing mechanism, is given by the following lemma. The lemma is formulated in terms of random variables (whose joint distribution is the CED) rather than in purely measure theoretic terms. This is not absolutely necessary, but it makes the formulation somewhat more transparent and intuitive, and closer in appearance to Definition 1.

Lemma 3. In a Bayesian game, a necessarily and sufficient condition for a correlated strategy distribution η to be a correlated equilibrium distribution is that some (equivalently, every[12]) pair of random variables $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \dots, \boldsymbol{t}_n)$ and $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ whose joint distribution is η satisfies

$$E(u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i}))|\boldsymbol{t}_i, \boldsymbol{a}_i) \geq 0, \qquad i \in N, a_i' \in A_i. \tag{11}$$

Proof. Let $\boldsymbol{m}$ be the canonical mechanism for η. The canonical correlated strategy σ with this mechanism is a correlated equilibrium if and only if it satisfies the condition in Definition 1. Since the corresponding random action profile $\boldsymbol{a}$ is the canonical one, i.e., $\boldsymbol{m}(\boldsymbol{t})$, that condition is equivalent to (11). This proves the sufficiency of the condition in the lemma, and it remains to prove its necessity.

---

[11] For a random variable $\boldsymbol{x}$ and a real-valued random variable $\boldsymbol{y}$, which are defined on the same probability space, the conditional expectation $E(\boldsymbol{y}|\boldsymbol{x})$ is also a random variable on that space. It is constant on every event of the form $[\boldsymbol{x} = x]$ (where $\boldsymbol{x}$ takes a particular value x), and its value there is $E(\boldsymbol{y}|\boldsymbol{x} = x)$, the conditional expectation of $\boldsymbol{y}$, given that $\boldsymbol{x} = x$. The meaning of equalities and inequalities involving conditional expectations is that they hold with probability 1 (equivalently, hold pointwise).

[12] The equivalence holds since whether inequality (11) below holds only depends on the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{a}$.

Every CED is by definition the joint distribution of a pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a}$ such that the latter is the random action profile corresponding to some correlated equilibrium σ with a mechanism $\boldsymbol{m}$ that satisfies the condition in Definition 1. For every player i, taking the condition expectation of both sides of (10), given $\boldsymbol{t}_i$ and $\boldsymbol{a}_i$, gives

$$E\left(E\left(u_i(\boldsymbol{t},\boldsymbol{a}) - u_i(\boldsymbol{t},(a_i',\boldsymbol{a}_{-i}))\middle|\boldsymbol{t}_i,\boldsymbol{m}_i(\boldsymbol{t})\right)\middle|\boldsymbol{t}_i,\boldsymbol{a}_i\right) \geq 0. \qquad (12)$$

Since, by (7), $\boldsymbol{a}_i$ can be expressed as a function of $\boldsymbol{t}_i$ and $\boldsymbol{m}_i(\boldsymbol{t})$, the iterated conditional expectation in (12) is equal to the (single) conditional expectation in (11) (see Shiryaev, 1996, Chapter I, §8). This proves that the condition in the lemma is also necessary for CED. ∎

A useful immediate corollary of Lemma 3 (which is already established in the first part of the proof) is that, to check whether a given CSD is a CED, it suffices to consider its canonical mechanism and strategy.

Corollary 1. A correlated strategy distribution is a correlated equilibrium distribution if and only if the canonical strategy, with the canonical mechanism, is a correlated equilibrium.

A probability measure η on $T \times A$ that satisfies the condition in Lemma 3 is sometimes referred to itself as a correlated equilibrium (Bergemann and Morris, 2007). Another reasonable alternative definition of this concept would be a mechanism with which the canonical strategy is a correlated equilibrium in the sense of Definition 1. Lemma 3 and Corollary 1 show that these two alternatives are not fundamentally different from the definition of correlated equilibrium given above. However, this paper emphatically distinguishes between correlated equilibrium and correlated equilibrium distribution, and between correlated equilibrium and the mechanism it uses. These distinctions are instrumental for the paper's primary objective of studying the implementability relation between a correlated equilibrium distribution and a mechanism, that is, the existence of some correlated equilibrium with the latter that gives the former.

2.5 Communication equilibria

Communication equilibrium differs from correlated equilibrium in that the players self-report their types to the mechanism. Correspondingly, the incentive-compatibility condition of correlated equilibrium is augmented by the requirement that a player cannot gain from being the only one to lie about his type (and, possibly, deviate from the correlated strategy). The reliance on the players' reports turns the mechanism from a primary source of information about the (other) players' types to a secondary source – a mediator. The mediator may be a physical entity, such as a disinterested third party or a piece of hardware, or it may be a communication protocol, such as a one-shot direct exchange of messages between two players.[13] Effectively, the message exchange in the last example is not limited to type reports. This is because each player could in principle use a gadget that takes a type

[13] The question of how, and to what extent, can unmediated communication between players replace a mediator or a correlation device lies outside to scope of this paper. This question has been extensively studied, for both complete and incomplete information games. See, for example, Forges (1990), Ben-Porath (2003), Gerardi (2004) and references therein.

as input and outputs the required message. The players' individual gadgets could then be viewed collectively as a single mechanism, in the sense of the definition in Section 2.2.

When a correlated strategy $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ with a mechanism $\boldsymbol{m}$ is used, a player i has an incentive to lie about his type if he can increase his expected payoff by misreporting it as some type t_i', thereby changing the (random) messages sent to the players to from $\boldsymbol{m}(t)$, where $t = (t_i, t_{-i})$ is the true type profile, to $\boldsymbol{m}(t_i', t_{-i})$. Player i may be able to take advantage of the resulting change of the other players' actions by altering the rule that determines his response to the messages the mechanism sends him, from σ_i to some $\sigma_i' : T_i \times M_i \to A_i$. The resulting random action profile $\boldsymbol{a}' = (\boldsymbol{a}_1', \boldsymbol{a}_2', \dots, \boldsymbol{a}_n')$ is given by

$$\boldsymbol{a}_i' = \sigma_i'(\boldsymbol{t}_i, \boldsymbol{m}_i(t_i', \boldsymbol{t}_{-i})), \qquad \boldsymbol{a}_j' = \sigma_j(\boldsymbol{t}_j, \boldsymbol{m}_j(t_i', \boldsymbol{t}_{-i})), \qquad j \in N \setminus \{i\}. \tag{13}$$

Communication equilibrium is defined by the requirement that, regardless of player i's true type, misreporting it in the above manner cannot increase the player's expected payoff.

Definition 2. In a Bayesian game, a correlated strategy σ with a mechanism $\boldsymbol{m}$ is a *communication equilibrium* if, for every player i, type t_i' for that player and function $\sigma_i' : T_i \times M_i \to A_i$,

$$E(u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, \boldsymbol{a}') | \boldsymbol{t}_i) \geq 0,$$

where $\boldsymbol{a}$ and $\boldsymbol{a}'$ are given by (7) and (13).

An extension of the revelation-principle argument used in the first paragraph of this subsection shows that the set of possible communication equilibrium outcomes would not change if players were allowed to send to the mechanism arbitrary (rather than just type) reports, possibly determined by private randomization. For a player i of type t_i, such a report would be described by a random variable $\boldsymbol{r}_i(t_i)$ with values in some finite (say) set R_i. Processing the reports would require a "generalized" mechanism, which assigns to each possible profile of reports $r = (r_1, r_2, \dots, r_n) \in R_1 \times R_2 \times \cdots \times R_n$ a random profile of messages $\boldsymbol{m}(r)$, which are sent back to the players. The action a_i for each player i would then be determined as a function $\hat{\sigma}_i$ by the player's type t_i, the report r_i he sent to the mechanism and the message m_i he got in response:

$$a_i = \hat{\sigma}_i(t_i, r_i, m_i).$$

The reason this setup is in fact no more general than the one described above is that there is a correlated strategy σ with a "normal" mechanism that produces identical actions. Essentially, the mechanism internalizes the players' reporting process. For each type profile t, the random message that the mechanism sends to each player i is the pair

$$(\boldsymbol{r}_i(t_i), \boldsymbol{m}_i(\boldsymbol{r}_1(t_1), \boldsymbol{r}_2(t_2), \dots, \boldsymbol{r}_n(t_n))).$$

The correlated strategy σ then determines i's action as a function of his type t_i and the message (r_i, m_i) he received according to

$$\sigma_i(t_i, (r_i, m_i)) = \hat{\sigma}_i(t_i, r_i, m_i).$$

It is a simple, and standard, exercise to show that if with the generalized mechanism none of the players i can increase his expected payoff by changing $(\boldsymbol{r}_i(t_i))_{t_i \in T_i}$ (which specifies the report player i sends) and/or $\hat{\sigma}_i$ (which specifies his response to the mechanism's messages), then the correlated strategy σ with the "normal" mechanism described above is a communication equilibrium in the sense of Definition 2.

A similar argument, which is also part of the revelation principle, shows that, in every Bayesian game, the set of communication equilibrium outcomes would not change also if the messages that the mechanism sends to the players were required to be concrete action recommendations rather than arbitrary objects (see Myerson, 1994). If a correlated strategy distribution is a *communication equilibrium distribution* (MED), that is, if it is the CSD of some communication equilibrium, then there is some mechanism in which the message spaces coincide with the respective players' actions spaces, and with which the canonical strategy of obeying the received message is a communication equilibrium that gives the MED. However, this does not mean that attention can be restricted to communication equilibria of this kind. As for correlated equilibria, such a restriction would be inconsistent with the goals of this paper, since it might affect in an unwarranted way the properties of the implementing mechanisms. For example, suppose that two players base their choice of actions on their own type and the type report that they receive from the other player. To implement this, it suffices to use a mechanism with property S, which simply transmits the reports. However, the same would not be true if the mechanism were also required to indicate each player's actual action. Since the actions depend on the players' own types, it would be impossible to maintain property S. More generally, the properties of the mechanism should only describe the properties of the communication channels available to the players (which may or may not involve a mediator). The way the players use these channels is expressed by a different entity, which is their correlated strategy.

3 Overview of Results

This section summarizes the main results in this paper and illustrates them by examples. The results are described in greater depth and detail in Sections 4, 5, 6 and 7.

3.1 Attributes of correlated strategy distributions

The six properties of mechanisms described in Section 2.2 are not independent. Property S implies $\tilde{S}$, O implies $\tilde{O}$, and D implies I. Therefore, for each of the three pairs, a mechanism may satisfy both properties, only the second one, or none of them. Altogether, there are $(3^3 =)$ 27 possibilities. This classification of mechanisms induces a classification of correlated strategy distributions. Each CSD has or does not have the attribute that it can be implemented by a mechanism with a particular property, or more generally a set of properties $\mathcal{P}$. For example, a CSD is S-*implementable* if it is implementable by some mechanism with property S, and it is S,O-*implementable*[14] if it is implementable by some mechanism that has both property S and property O. The various attributes of CSDs are not independent. For example, S,O-implementability implies $\tilde{S}$,O-implementability, since S is a more stringent requirement than $\tilde{S}$, and it is implied by S,O,I-implementability, which involves the additional requirement that the implementing mechanism also satisfies I. A

[14] Since it is the *set* of properties that matters, and not their order, $\{S,O\}$-implementability might be a better notation. However, for the sake of readability, the curly brackets are omitted.

natural question, for each of these implications, is whether the reverse implication also holds, so that the two attributes are actually equivalent. As the Hesse diagram in Figure 1 shows, the answer is affirmative for S,O- and $\tilde{S},O$-implementability (which are equivalent) but negative for S,O- and S,O,I-implementability (which are not equivalent). Thus, every CSD that is implementable by some mechanism that satisfies $\tilde{S}$ and O is also implementable by a mechanism that satisfies S and O (and vice versa), and there is some such CSD in some Bayesian game that is not implementable by any mechanism that also satisfies I.

As Figure 1 shows, there are not 27 but only 7 distinct (i.e., nonequivalent) attributes of CSDs that can be defined in terms of the properties of the implementing mechanisms. Each can be described in several equivalent ways by using different combinations of properties. For example, S-implementability and $\tilde{S}$-implentability are both equivalent to the attribute of simply being a CSD, which is denoted in Figure 1 by the empty set (of properties of implementing mechanisms) $\{\,\}$. Thus, the limitations that these two properties put on the implementing mechanisms are inconsequential. Note that the seven attributes of CSDs are not all comparable. That is, some attributes do not imply and are not implied by certain other attributes.

Additional attributes of correlated strategy distributions in Bayesian games may conceivably be defined by conjunction. For example, a CSD may be both $\tilde{O}$-implementable *and* D-implementable. A natural question is whether this is equivalent to $\tilde{O},D$-implementability. More generally, if there is some implementing mechanism with a particular set of properties and another mechanism with some other properties, does it follow that the CSD is implementable by a single mechanism that has all the properties of the other two? Lemma 5 in Section 4 answers this question in the affirmative. An immediate corollary of this result (see Section 3.4) is that conjunctions do not in fact define new attributes of CSDs.

3.1.1 Intrinsic characterizations

Each of the seven attributes of CSDs in Figure 1 can also be characterized *intrinsically*, that is, without explicitly referring to implementing mechanisms. The significance of intrinsic characterizations is that they may make it easier to check whether a particular distribution has a particular attribute. Lemma 2 may be viewed as an intrinsic characterization of the weakest attribute, which is simply being a CSD (I in Figure 1). The following proposition characterizes the strongest attribute, which is being a pure-strategy distribution (VII in Figure 1), as well as the attribute of being a mixed-strategy distribution (V in Figure 1).

Proposition 1. A correlated strategy distribution in a Bayesian game is a mixed-strategy distribution if and only if it is the joint distribution of a pair of random variables $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \dots, \boldsymbol{t}_n)$ and $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ such that

(i) For each player i, $\boldsymbol{a}_i$ and $(\boldsymbol{t}_{-i}, \boldsymbol{a}_{-i})$ are conditionally independent, given $\boldsymbol{t}_i$.

A correlated strategy distribution is a pure-strategy distribution if and only if it satisfies the stronger condition in which (i) is replaced by

(ii) For each player i, the conditional distribution of i's action $\boldsymbol{a}_i$, given his type $\boldsymbol{t}_i$, is degenerate.

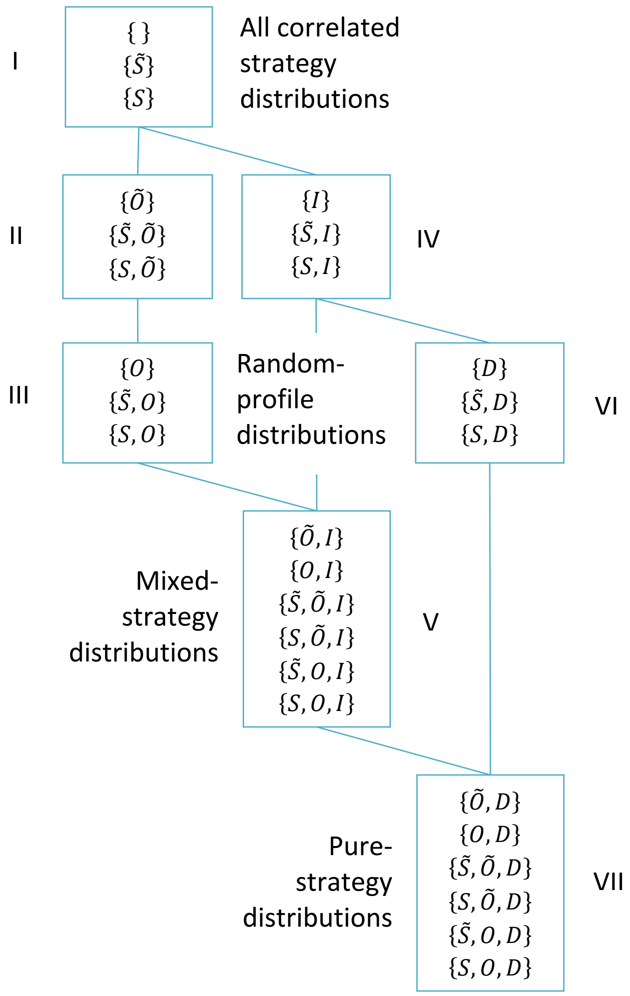

Figure 1. Hesse diagram of the different attributes of correlated strategy distributions (CSDs) in Bayesian games, ordered by implication. An attribute is represented by a box containing its equivalent definitions, each of which is a set of properties possessed by some mechanism that implements the CSD. Two sets in the same box identify mechanisms that implement exactly the same CSDs. For those in different boxes, the implementable CSDs are different. A line segment connecting two boxes indicates that the lower attribute implies the upper one but the reverse implication does not hold.

In other words, pure-strategy distributions are characterized by the property that if a player's type is known, there is no uncertainty about his action. Mixed-strategy distributions are characterized by the property that if a player's type is known, his action does not add any information about the other players' types or actions.[15] Proposition 1 is proved in Section 4.

The next proposition characterizes random-profile distributions (III in Figure 1). The characterizing property is the existence of a probability measure μ on $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ that satisfies a certain condition. Such a measure assigns a probability to each pure-strategy profile $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$ (see Section 2.1). For each type profile $t = (t_1, t_2, \dots, t_n)$ there is a corresponding marginal measure μ^t on $A_1 \times A_2 \times \cdots \times A_n$, which assigns to each action profile $a = (a_1, a_2, \dots, a_n)$ the probability that the actions

[15] For an extension of this result to games with a random number of players, see Milchtaich (2004, Theorem 2).

associated with the players' types specified by t are those specified by a. Formally,

$$\mu^t(\{a\}) = \mu\big(\{(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots) \in A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n} \,|\, (a_1^{t_1}, a_2^{t_2}, \dots, a_n^{t_n}) = a\}\big). \tag{14}$$

The proof of the following proposition is given in Section 4, and it is illustrated by Example 1 below.

Proposition 2. A correlated strategy distribution η is a random-profile distribution if and only if there is a probability measure μ on $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ such that

$$\eta(\{(t,a)\}) = \eta_T(\{t\})\, \mu^t(\{a\}), \quad t \in T, a \in A,$$

where η_T is the distribution of type profiles and μ^t is the marginal measure defined by (14).

A random-profile distribution is not necessarily a mixed-strategy distribution. Therefore, by Proposition 1, it may not have the property that if a player's type is known, his action does not add any information about the other players' types or actions. However, a random-profile distribution always has the weaker property that, if a player's type is known, his action does not add any information about the other players' types (but may say something about their actions). This property can be expressed formally as follows (Forges, 1993).

Definition 3. A correlated strategy distribution η has the *conditional independence property* if for some (equivalently, every) pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a}$ whose joint distribution is η, the action $\boldsymbol{a}_i$ of each player i and the types $\boldsymbol{t}_{-i}$ of the other players are conditionally independent, given i's own type $\boldsymbol{t}_i$.

The conditional independence property is not, however, unique to random-profile distributions. As shown by the following example (which is based on Example 6 in Milchtaich, 2004; see also Lehrer at al., 2010, and Forges, 2006), even for two-player games the two conditions are not equivalent.[16] Whereas being a random-profile distribution (which by definition means S, O-implementability) is equivalent to O-implementability, it is shown by Proposition 3 in Section 4 that the conditional independence property is equivalent to the weaker attribute of $\tilde{O}$-implementability (equivalently, $S, \tilde{O}$-implementability; II in Figure 1).

Example 1. *A correlated strategy distribution with the conditional independence property that is not a random-profile distribution.* The two players in a 2×2 Bayesian game have identical action spaces, $A_1 = A_2 = \{L, R\}$, and identical two-element type spaces, $T_1 = T_2 = \{+1, -1\}$. The four type profiles are equally probable, so that types are independent. (Independence is not a crucial assumption. It would suffice to assume that all type profiles have positive probability.) A correlated strategy distribution is defined as follows: (i) If both players have type $+1$, the action profile is either (L, L) or (R, R), each

[16] Forges (1993) and Cotter's (1994) suggestion that the two properties are equivalent is mistaken. The mistake was corrected in Forges (2006). As explained below, the non-equivalence reflects the difference between properties O and $\tilde{O}$ of devices.

with probability 0.5, and (ii) if the type profile is any of the other three, the action profile is either (L, R) or (R, L), each with probability 0.5.

This correlated strategy distribution has the conditional independence property, since for each type profile, each player takes action L with probability 0.5. However, this is not a random-profile distribution. This can be proved by assuming that a measure μ on pure-strategy profiles as in Proposition 2 exists, and showing that this assumption leads to a contradiction. Since, by (ii), there is zero probability that the players take identical actions when the type profile is different from $(+1, +1)$, μ must, in particular, assign zero probability of any pure-strategy profile of the form $(R,*;*, R)$, $(*, R; R,*)$ or $(*, L;*, L)$, where a wildcard action $*$ can be either L or R. The same must therefore be true for any pure-strategy profile of the form $(R,*; R,*)$, which necessarily also has one of the above three forms. However, this implies that there is zero probability that both players play R when the type profile is $(+1, +1)$, which contradicts (i).

Intrinsic characterizations for the remaining two attributes of CSDs (IV and VI in Figure 1) are given by Proposition 4 in Section 4.

3.2 Attributes of correlated equilibrium distributions

A correlated equilibrium distribution in a Bayesian game is also a correlated strategy distribution but the converse is not always true. However, since every correlated strategy distribution can be made a correlated equilibrium distribution simply by replacing the payoff functions by constant ones, the number of distinct (i.e., nonequivalent) attributes of CEDs that can be defined in terms of the properties of the implementing mechanisms is not smaller than for CSDs. In fact, as Figure 2 shows, the number is significantly larger: 14 or 15 instead of 7. This reflects the fact that the classifications of CEDs can be viewed as consisting of two layers: (i) the classification induced by the different attributes of CSDs, and (ii) the refinement that results from also taking into account the incentive compatibility requirement (expressed by Lemma 3). Thus, two CEDs in a Bayesian game may differ (i) in that even as CSDs they require different kinds of implementing mechanisms, or (ii) only in that different kinds of implementing mechanisms are compatible with the equilibrium condition. This is a useful distinction, which seems to be lacking in the existing literature on games with incomplete information.

Where the equilibrium condition is effective is the connection between a player's type and the messages he receives, i.e., properties S and $\tilde{S}$ of the mechanism. For CSDs these properties do not make any difference, as can be seen in Figure 1, but this is not so for CEDs. For example, as can be seen in Figure 2, for CEDs the three attributes of simply being a CED, S-implementability and $\tilde{S}$-implementability (I, I_a and I_b in Figure 2) are not equivalent. Thus, there are CEDs in some Bayesian games that cannot be implemented by any mechanism satisfying $\tilde{S}$, and there are CEDs that are implementable by such a mechanism but cannot be implemented by any mechanism with the stronger property S. The following two examples present such CEDs. Note that these examples, like the other ones in this subsection and Example 1, concern two-player games. Therefore, the Hesse diagrams in Figure 1 (CSDs) and Figure 2 (CEDs) apply to two-player Bayesian games as well as to the general, n-player case.

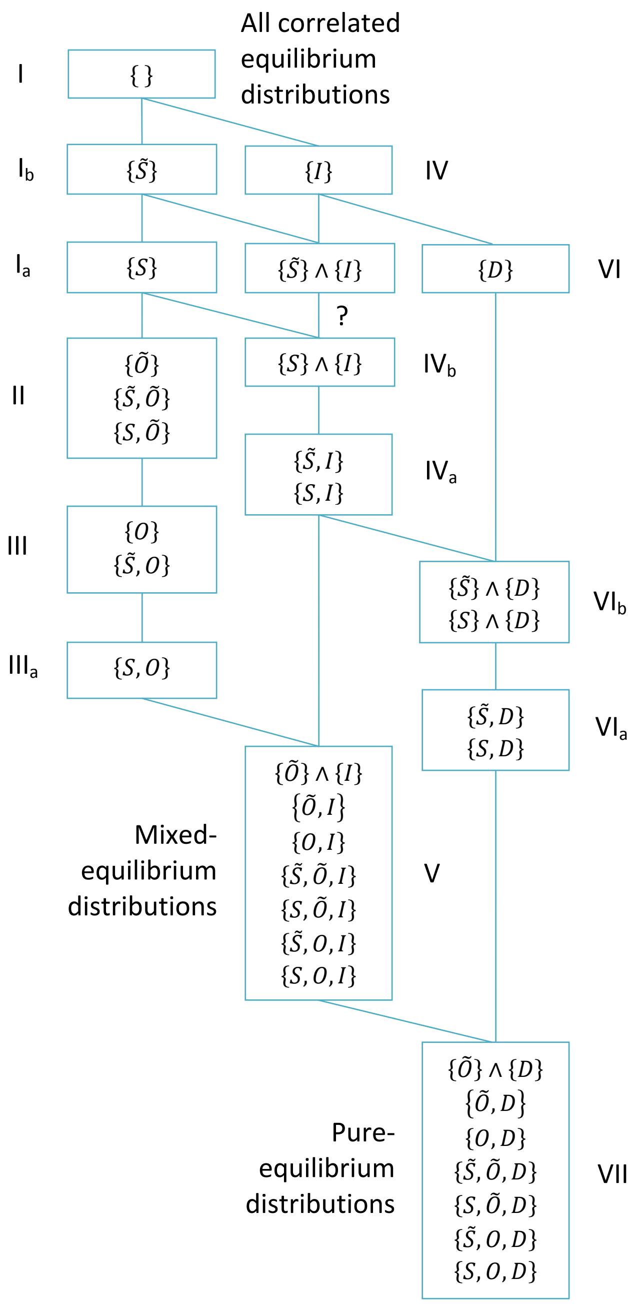

Figure 2. Hesse diagram of the different attributes of correlated equilibrium distributions (CEDs) in Bayesian games, ordered by implication. As in Figure 1, each attribute is represented by a box, and an implication relation is represented by a line. A conjunction symbol ∧ means that the CED is implementable both by a mechanism with one property and by a mechanism satisfying the other property. The line marked ? represents an uncertain relation: it is not known whether the reverse implication also holds (in which case the two connected boxes should be coalesced).

Example 2. *A correlated equilibrium distribution that is not $\tilde{S}$-implementable.* In a symmetric 2×2 game with the game structure and distribution of type profiles described in Example 1, the two players always get equal payoffs, which for a type profile (t_1, t_2) are given by the payoff matrix

$$\begin{matrix} & L & R \\ \begin{matrix} L \\ R \end{matrix} & \multicolumn{2}{c}{\begin{pmatrix} -t_1 t_2 & 0 \\ 0 & 2t_1 t_2 \end{pmatrix}} \end{matrix}.$$

Thus, depending on the type profile, L or R is a dominant action for both players. The message that each player receives is a type profile. For both types of player 1 and for type $+1$ of player 2, this type profile is the real one (t_1, t_2). However, for type -1 of player 2, the message is always $(-1, -1)$. With this mechanism, the correlated strategy that instructs each player to choose the dominant action for the type profile specified by the message he receives is a correlated equilibrium. Types $+1$ and -1 of player 1 and type $+1$ of player 2 clearly cannot do any better than playing according to the strategy, which gives them the expected payoffs 1, $3/2$ and $3/2$, respectively. The same is true for type -1 of player 2, whose expected payoff would decrease from 1 to $1/2$ if he switched from his (constant) action R to L.

The mechanism described above does not involve randomization, and thus has property D, but it does not satisfy $\tilde{S}$. In fact, the corresponding correlated equilibrium distribution is not implementable by any mechanism with the latter property. Specifically, it is not implementable by any mechanism for which, for each type of player 1, the message that player 2 receives is (statistically) independent of his own type. To see this, suppose that such a mechanism exists, and let m_2^+ and m_2^- be some specific messages that player 2 receives with positive probability (which is the same for both types of that player) when player 1's type is $+1$ and -1, respectively. The two messages cannot be identical, for otherwise type $+1$ of player 2 would receive this message with positive probability both when 1's type is $+1$ and when the type if -1, which is inconsistent with the fact that (according to the above distribution) with probability 1 he plays R in the former case and L in the latter. Therefore, every such m_2^+ and m_2^- must be distinct, which implies that player 2 can always tell by his message the type, and hence also the action, of his opponent. However, this is inconsistent with incentive compatibility, since it implies that, by choosing the same action as the opponent, type -1 of player 2 could increase his payoff from 1 to $3/2$. This contradiction proves that the above correlated equilibrium distribution is not implementable by any mechanism satisfying $\tilde{S}$.

Example 3. *A correlated equilibrium distribution that is $\tilde{S}$- but not S-implementable.* The game is the same as in Example 2. A correlated equilibrium distribution in this game is given by Table 1, which specifies the (conditional) distribution of the players' action profile for each type profile. As seen in the table, the marginal distributions, i.e., the probability that a player of a given type plays L or R, depend only on the opponent's type. (Specifically, L has probability 0.75 or 0.5 if the type of the opponent is $+1$ or -1, respectively.) Therefore, a mechanism that randomly chooses an action profile (a_1, a_2) according to the probabilities corresponding to the players' actual type profile and reports a_1 to player 1 and a_2 to player 2 has property $\tilde{S}$. The correlated strategy of acting according to the messages is a correlated equilibrium; it is not difficult to check that a player can never increase his expected payoff by deviating to the other action. It is also true, but less easy to check, that the CED in Table 1 is not implementable by any mechanism that has the stronger property S. In fact, it takes a computer to check this. Although the problem is a standard linear programming one – it needs to be checked that a particular system of linear equalities and inequalities does not have a solution – the number of variables and equalities/inequalities involved (at least 256 and 20, respectively) is far too great for manual calculations.

			Player 2						
			Type $+1$				Type -1		
			L	R			L	R	
Player 1	Type $+1$	L	0.75	0	0.75	U	0.25	0.25	0.5
		R	0	0.25	0.25	D	0.5	0	0.5
			0.75	0.25			0.75	0.25	
			L	R			L	R	
	Type -1	L	0.25	0.5	0.75	U	0	0. 5	0.5
		R	0.25	0	0.25	D	0.5	0	0.5
			0.5	0.5			0.5	0.5	

Table 1. A correlated equilibrium distribution for Example 3. The four type profiles are equally probable. For each of them, the joint distribution of player 1's and player 2's actions (which can be L or R), as well as the marginal distributions, are given.

Some intuition[17] about why a mechanism with property S cannot implement the CED in Table 1 can be gained by considering two conceivable mechanisms with this property. The first mechanism generates the messages by randomizing over pure-strategy profiles according to a suitable distribution and then recommending to each player the action that his pure strategy specifies for the *opponent's* actual type. In this way, a player's own type does not affect the message he receives. This mechanism fails on a very basic level: it cannot implement the distribution in Table 1 even as a correlated strategy distribution. The problem is similar to that in Example 1, and it does not involve incentives (i.e., payoffs).

The second conceivable mechanism for the CED in Table 1 has property S by virtue of sending as a message to each player not a single action but a pure strategy, and leaving it to the player to choose the action corresponding to his actual type. (Such a mechanism is used in the proof of Proposition 5 below.) Specifically, the mechanism first chooses action profiles according to the probabilities specified in Table 1, independently for each of the four type profiles. Then, based on the players' actual type profile (t_1, t_2), it tells player 1 both his action for the type profile $(+1, t_2)$ and his action for $(-1, t_2)$, and similarly for player 2. If the players use the messages in the intended manner, i.e., take the first or second action if their type is $+1$ or -1, respectively, then the action distributions for the four type profiles are indeed as in Table 1. However, this correlated strategy is not a correlated equilibrium. The reason is that the double message conveys too much information about the opponent's type. By assumption, the prior probability that player 2 has type $+1$ is 0.5. By Bayes' rule, the posterior probability that 2 has that type given that player 1 plays L is 0.6. Thus, telling player 1 which action he should take gives him some information about 2's type, but not too much information, in that taking the action is still optimal for him. (As indicated in the previous paragraph, telling the players *only* the actions they should take cannot be implemented by a mechanism that satisfies S.) A double message as above amounts to two independent draws from the same unknown distribution, which are more informative about the underlying distribution than a single draw. For example, telling player 1 that he should play L whether his type is $+1$ or -1 increases the (posterior) probability that 2's type is $+1$ to almost 0.7. This probability is greater than 2/3, which implies that, regardless of the actions the two types of play 2 take, R is the better action for type $+1$ of player 1. Thus, a player may deduce from the additional information conveyed by the double message that

[17] Note that this is not meant to be an outline of a proof.

his expected payoff from taking the action he is supposed to take is actually less than for the alternative action.

Examples 2 and 3 illustrate a point of general significance. The reason S and $\tilde{S}$ affect implementability of CEDs is that these properties may entail inability to restrict messages to certain types of players only. This is not a problem for correlated strategies, where information cannot do any harm, but it may be a problem for correlated equilibria, where incentive compatibility matters. However, as can be seen in Figure 2, whether this is actually so depends on the other properties of the mechanism. For example, for $\tilde{O}$-implementable CEDs (II in Figure 2), requiring that the implementing mechanism also satisfy S or $\tilde{S}$ does not make any difference.

3.2.1 Attributes inherited from correlated strategy distributions

Since a correlated equilibrium distribution in a Bayesian game is in particular a correlated strategy distribution, it has as such one or more of the attributes in Figure 1. However, a CED that as a CSD has the attribute that it is implementable by a mechanism with a particular set of properties is not necessarily implementable by such a mechanism as a CED. That is, it may be impossible to find a suitable correlated strategy with that kind of mechanism that is also a correlated equilibrium. For example, the CED in Example 2 is not $\tilde{S}$-implementable even though it is $\tilde{S}$-implementable as a CSD (since this is so for every CSD; see Figure 1). However, as the following theorem shows, this kind of discrepancy between the two notions of implementability only arises when properties S or $\tilde{S}$ of mechanisms are involved. Since the other four properties of mechanisms defined in Section 2.2 are sufficient to characterize all the attributes of CSDs in Figure 1, this means that a CED has attribute I, II, III, IV, V, VI or VII in Figure 2 if and only if, as a CSD, it has the similarly numbered attribute in Figure 1. Thus, for example, a CED is a pure- or mixed-equilibrium distribution if and only if is a pure- or mixed-strategy distribution, respectively. The proof of the theorem is given in Section 5.

Theorem 1. For $\mathcal{P} \subseteq \{O, \tilde{O}, D, I\}$, a correlated equilibrium distribution is implementable by a mechanism with all the properties in $\mathcal{P}$ if and only if it satisfies a similar condition as a correlated strategy distribution.

It follows from Theorem 1 that the intrinsic characterizations for the seven attributes of CSDs given in Section 3.1.1 and (Propositions 3 and 4 in) Section 4 also apply to the corresponding attributes of CEDs. For example, a CED is $\tilde{O}$-implementable (attribute II in Figure 2) if and only if it has the conditional independence property. An intrinsic characterization for the very attribute of being a CED (I in Figure 2) is given by Lemma 3, which says that a CSD is a CED if and only if it satisfies a certain incentive-compatibility condition (for distributions). It follows that, for example, an $\tilde{O}$-implementable CED can be (intrinsically) characterized as a CSD that satisfies the incentive-compatibility condition and has the conditional independence property.

3.2.2 Attributes defined by conjunction

A significant difference between attributes of CSDs and those of CEDs concerns the effect of conjunctions. As indicated in Section 3.1, if a CSD that is implementable by a mechanism with one of the properties in Section 2.2 is also implementable by a mechanism with another property, then it is implementable by a single mechanism that has both properties. It follows from Theorem 1 that this remains true for CEDs as long as the properties concerned are not S or $\tilde{S}$. However, as the following two examples show, this is not so in general. The first example presents a CED that is implementable by a mechanism with property S and by a mechanism with property D but is not implementable by any mechanism with both properties. The second example replaces D with I.

Example 4. *A correlated equilibrium distribution that is S- and as well as D-implementable but not S, D-implementable.* In a two-player Bayesian game, player 1 has two types, t_1' and t_1'', and two actions, L and R. Player 2 has three types, t_2', t_2'' and t_2''', and only one action, L. All type profiles except (t_1', t_2') may occur, and they have the same probability (1/5). If player 1 plays R, the payoff to both players is 0.5. If he plays L, the payoff vector is determined by the type profile according to the following table:

	t_2'	t_2''	t_2'''
t_1'	N/A	(1,0)	(0,1)
t_1''	(3,0)	(0,0)	(0,0)

The lowest possible expected payoff for player 2 in this game is 0.1. As shown below, there is a unique correlated equilibrium distribution with this payoff, and this CED is both D-implementable and S, I-implementable but it is not even $\tilde{S}, D$-implementable.

For player 2's expected payoff to be 0.1, player 1 should play R if and only if the type profile is (t_1', t_2'''). Consider the implementing mechanism with property D (i.e., no randomization) that sends to player 1 the message R if the type profile is (t_1', t_2''') and otherwise sends L. The correlated strategy in which player 1 follows the mechanism's instructions is a correlated equilibrium, since it always gives maximum payoff to type t_1' of player 1 and gives t_1'' (who is always instructed to play L) an expected payoff of 1, which is greater than the 0.5 he would receive from playing R. Another mechanism, with properties S and I, that implements the same CED is a mechanism that sends to player 1 the message L or R if player 2 has type t_2'' or t_2''' respectively, and sends either message with probability 1/2 if the type is t_2'. The correlated strategy in which player 1 follows the mechanism's instructions if his type is t_1' but always plays L if the type is t_1'' is a correlated equilibrium, with the same CED. This is because the message that type t_1'' of player 1 receives does not affect the probability that he assigns to player 2's type being t_2', which is 1/3 regardless of the received message.

It remains to show that any mechanism with properties $\tilde{S}$ and D cannot implement the above CED. The message that player 1 receives from a mechanism with these properties must be a function of player 2's type, say m_1', m_1'' or m_1''' if the type is t_2', t_2'' and t_2''', respectively. To implement the CED, in which the action that type t_1' of player 1 takes depends on whether or not 2's type is t_2''', the message m_1'' must be different from m_1'''. Therefore, one of these, say m_1'', must also be different from m_1'. But this means that the mechanism effectively tells player 1 when 2's type is t_2''. Therefore, in any correlated equilibrium with that mechanism, player 1 plays (his payoff-maximizing action) R, and not L, when the type profile is (t_1'', t_2'').

		Player 2						
		Type $+1$				Type -1		
Player 1		L	R			L	R	
Type $+1$	L	2	0	2/3	L	0	0	2/3
	R	0	1	1/3	R	1	0	1/3
		1/2	1/2			1/2	1/2	
		L	R			L	R	
Type -1	L	0	0	1/3	L	0	0	1/2
	R	0	0	2/3	R	0	0	1/2
		1/2	1/2			1/2	1/2	

Table 2. A correlated equilibrium distribution for Example 5. The four type profiles are equally probable. For each of them, the actions that players 1 and 2 take are independent. The probabilities that these actions are L or R are given at the margins of the corresponding box. The numbers inside the box are player 1's payoffs. Player 2 always gets payoff 0.

Example 5. *A correlated equilibrium distribution that is S- as well as I-implementable but not S, I-implementable.* The game structure and distribution of type profiles are again as in Examples 1, 2 and 3. The payoff matrices of player 1, one for each type profile, are shown in Table 2. Player 2 has the constant payoff function 0. A mechanism with property I randomly chooses an action for each player for each type profile according to the probabilities shown in Table 2, such that these eight choices are independent. It then tells each player the action chosen for him for the actual type profile. The players' strategy is to play accordingly. This is a correlated equilibrium. The reason is that a change of action by player 1 may affect his payoff only if his type is $+1$ and (i) player 2 has type $+1$ and he plays L, (ii) player 2 has type $+1$ and he plays R, or (iii) player 2 has type -1 and he plays L. The effect in case (i) has the opposite sign and twice the magnitude of the effect in the other two cases. Since (i), (ii) and (iii) always has equal conditional probabilities, given that the type of player 1 is $+1$ and given his action, this means that the conditional expectation of the gain from changing action is always zero.

The same CED is implementable by a mechanism with property S. The mechanism first chooses two pairs of actions, $a^+ = (a^{++}, a^{-+})$ and $a^- = (a^{+-}, a^{--})$. The pairs are chosen independently of one another, the probability that a^+ equals (L, L), (L, R), (R, L) or (R, R) is $1/6$, $1/2$, $1/6$ and $1/6$, respectively, and for a^- the corresponding probabilities are $1/3$, $1/3$, $1/6$ and $1/6$. Then, for each type profile $t = (t_1, t_2)$, the mechanism chooses an action a_2^t for player 2, with probabilities (for R and L) that depend on (both t and) a^- (that was chosen in the first stage). Specifically, the probability that $a_2^t = L$ is $1/2$ *unless* $t = (+1, -1)$ and (i) $a^- = (L, L)$, in which case the probability is $1/4$, or (ii) $a^- = (L, R)$, in which case the probability is $3/4$. Finally, the mechanism sends messages to the players, which depend on the choices made in the first two stages and on the players' actual type profile $t = (t_1, t_2)$. The message to player 1 is a^+ or a^- if 2's type is $+1$ or -1, respectively, and the message to player 2 is the pair of actions $(a_2^{(t_1,+1)}, a_2^{(t_1,-1)})$. Thus, neither message reflects the player's own type. It is not very difficult to check that the correlated strategy specifying that each player chooses the first or second action in his message if his type is $+1$ or -1, respectively, gives the CED described above. For example, if $t = (+1, +1)$, the action profile is $(a^{++}, a_2^{(+1,+1)})$, which is (L, R), (L, R), (R, L) or (R, R) with probability $1/3$, $1/3$, $1/6$ and $1/6$, respectively. Therefore, the players' actions are independent and are distributed as specified at the margins of the top-left box in Table 2.

To show that the above correlated strategy (with the described mechanism with property S) is a correlated equilibrium, it suffices to prove that, given that the type of player 1 is $+1$ and given the message he receives (which can be (L, L), (L, R), (R, L) or (R, R)), the conditional probabilities of the following three events are equal: (i) player 2 has type $+1$ and he plays L, (ii) player 2 has type $+1$ and he plays R, and (iii) player 2 has type -1 and he plays L. As indicated above, such equality means that player 1 is indifferent between his two actions. The equality can be viewed as the conjunction of two equalities: (a) events (i) and (ii) have equal conditional probabilities, which are necessarily one-half the conditional probability that $t_2 = +1$, and (b) the latter is also equal to twice the conditional probability of (iii). To prove (a) it suffices to note that, given that $t = (+1, +1)$, the message a^+ that player 1 receives and the action $a_2^{(+1,+1)}$ that player 2 takes are conditionally independent, and the probability that the latter is L is $1/2$. To prove (b), note, first, that by the specification of the mechanism and Bayes' rule the conditional probability that $t_2 = +1$, given that player 1's type is $+1$ and he receives the message (L, L), (L, R), (R, L) or (R, R), is $1/3$, $3/5$, $1/2$ or $1/2$, respectively. It is therefore sufficient to show that the conditional probability, given the same information, that $t_2 = -1$ and $a_2^{(+1,-1)} = L$ is $1/6$, $3/10$, $1/4$ or $1/4$, respectively. This conditional probability is equal to the product of two terms: the condition probability that $t_2 = -1$, given that $t_1 = +1$ and player 1's message has the specified value; and the condition probability that $a_2^{(+1,-1)} = L$, given that $t = (+1, -1)$ and a^- has that value. The first term is the complement of the conditional probability that $t_2 = +1$, and is hence $2/3$, $2/5$, $1/2$ or $1/2$ if the message is (L, L), (L, R), (R, L) or (R, R), respectively; and by the specification of the mechanism, the second term is $1/4$, $3/4$, $1/2$ or $1/2$, respectively. Therefore, the product of the two terms is $1/6$, $3/10$, $1/4$ or $1/4$, respectively, as had to be shown.

The above CED, which as shown is both S- and I-implementable, is not implementable by any mechanism that has both properties, or even by a mechanism with properties $\tilde{S}$ and I. To see this, consider any correlated strategy with a mechanism satisfying $\tilde{S}$ and I that has the CSD specified by Table 2. Partition all the messages that player 1 may receive into four groups, (L, L), (L, R), (R, L) and (R, R), according to the actions that player 1's strategy assigns to the message when the player's type is $+1$ (first entry) and -1 (second entry). Since the mechanism satisfies $\tilde{S}$, the probability of receiving a message that belongs to a particular group when player 2 has type $+1$ is the same for both types of player 1. Denote these probabilities by p_{LL}^+, p_{LR}^+, p_{RL}^+ and p_{RR}^+. Let p_{LL}^-, p_{LR}^-, p_{RL}^- and p_{RR}^- be the corresponding probabilities for the case where player 2's type is -1. Since these messages induce the distributions of actions given in Table 2, the following equalities must hold:

$$p_{LL}^+ + p_{LR}^+ = \frac{2}{3}\,, \qquad p_{LL}^- + p_{LR}^- = \frac{2}{3}, \tag{15}$$

$$p_{LL}^+ + p_{RL}^+ = \frac{1}{3}\,, \qquad p_{LL}^- + p_{RL}^- = \frac{1}{2}. \tag{16}$$

A necessary condition for the correlated strategy to be a correlated equilibrium is that type $+1$ of player 1 cannot increase the conditional expectation of his payoff by playing R, R, L or L, respectively, when the message he receives belongs to group (L, L), (L, R), (R, L) or

(R, R) (so that his strategy specifies taking the opposite action). Since the mechanism satisfies I, for any type profile the message that player 1 receives is independent of player 2's message, and hence of that player's action. Thus, regardless of player 1's type and the message he receives, player 2 plays L with probability $1/2$. The above necessary condition is therefore expressed by the following inequalities:

$$
\begin{aligned}
&p_{LL}^{+}\left(-\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1\right)+p_{LL}^{-}\left(\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 0\right)\leq 0,\\
&p_{LR}^{+}\left(-\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1\right)+p_{LR}^{-}\left(\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 0\right)\leq 0,\\
&p_{RL}^{+}\left(\frac{1}{2}\cdot 2-\frac{1}{2}\cdot 1\right)+p_{RL}^{-}\left(-\frac{1}{2}\cdot 1-\frac{1}{2}\cdot 0\right)\leq 0,\\
&p_{RR}^{+}\left(\frac{1}{2}\cdot 2-\frac{1}{2}\cdot 1\right)+p_{RR}^{-}\left(-\frac{1}{2}\cdot 1-\frac{1}{2}\cdot 0\right)\leq 0.
\end{aligned}
\tag{17}
$$

All inequalities in (17) must in fact hold as equalities. If one of the first two inequalities or one of the last two were strict, then $-(p_{LL}^{+}+p_{LR}^{+})+(p_{LL}^{-}+p_{LR}^{-})<0$ or $(p_{RL}^{+}+p_{RR}^{+})-(p_{RL}^{-}+p_{LR}^{-})<0$ would hold. These two inequalities are equivalent (since the probabilities in each quartet sum up to 1), and they contradict (15). Therefore, in particular, the first and third equalities in (17) hold as equalities, which implies $-(p_{LL}^{+}+p_{RL}^{+})+(p_{LL}^{-}+p_{RL}^{-})=0$. This equation contradicts (16). The contradiction proves that a correlated strategy with a mechanism satisfying $\tilde{S}$ and I that has the distribution specified by Table 2 cannot be a correlated equilibrium.

The conjunction of S- and D-implementability (as in Example 4) and the conjunction of S- and I-implementability (Example 5) are two attributes of CEDs that have no parallels among the attributes of CSDs. A third attribute that is defined in a similar manner may exist, namely, the conjunction of $\tilde{S}$- and I-implementability. However, its existence is still an open question: it is not known whether this third attribute is indeed distinct from the second one. This uncertainty is represented in Figure 2 by the question mark. It is shown in Section 5.3 below that, in any case, these two or three attributes of CEDs are the *only* ones that can be defined only by conjunctions; additional such attributes do not exist.

3.2.3 Strategy correlated and type correlated equilibria

As an illustration of the discussion in the previous subsections, this subsection describes in detail two of the attributes of correlated equilibrium distributions in Figure 2: O-implementability (attribute III), which is inherited from correlated strategy distributions, and the "spin-off" attribute S, O-implementability (attribute III_a). Both attributes have been previously described in the literature, under various names. Correlated equilibria whose CEDs are S, O-implementable are called strategy correlated equilibria (Cotter, 1991), strategic form correlated equilibria (Forges, 1993, 2006) or normal form correlated equilibria (Lehrer et al., 2010). Correlated equilibria whose CEDs are O-implementable are called type correlated equilibria (Cotter, 1994) or agent normal form correlated equilibria (Forges, 1993, 2006; Lehrer et al., 2006, 2010).

In a *strategy correlated equilibrium*, a referee who does not know the players' types confidentially recommends a strategy to each of them. The recommendations are thus independent of the players' actual types but not necessarily of one another. The equilibrium condition is that it is always optimal for each player to take the action that the strategy recommended by the referee prescribes to his actual type, assuming that all the other players do the same.[18] In the terminology of this work, a referee corresponds to a mechanism. The assumption that the referee does not know the players' types corresponds to properties S and O of the mechanism, which together mean that the types do not affect the messages.

A *type correlated equilibrium* differs from a strategy correlated equilibrium in that each player is told not the whole strategy but only the action it prescribes to his actual type. However, it is still assumed that the referee does not know the types when he chooses his recommendations. Hence, either he learns them later or there is something (e.g., a language barrier; see the example in the Introduction) that prevents players from learning what they are not supposed to know, namely, the actions that the strategy recommended by the referee prescribes to each of the other types of the same player. Either way, the message that each player receive may depend on his type, so that the mechanism only has property O.

Forges (1993) showed that some type correlated equilibria are not equivalent to any strategy correlated equilibrium. The next example provides another demonstration of this result.

Example 6. *A correlated equilibrium distribution that is $\tilde{S}, O$- but not S, O-implementable.* Two players play the coordination game

$$\begin{matrix} & L & R \\ L & 1 & 0 \\ R & 0 & 1 \end{matrix}.$$

Player 1 can be of type $+1$ or type -1, which are both equally likely. Player 2 has the single type $+1$. A mechanism bases its messages on the outcomes of two independent coin tosses, $\boldsymbol{m}^+$ and $\boldsymbol{m}^-$, each of which gives L or R with equal probabilities. A player of type $+1$ or -1 receives the message $\boldsymbol{m}^+$ or $\boldsymbol{m}^-$, respectively. The correlated strategy of always acting according to the message is a correlated equilibrium. It gives the expected payoffs 1 and 0.5, respectively, to types $+1$ and -1 of player 1, and 0.75 to player 2, and it is easy to check that, in all three cases, profitable deviations do not exist.

The above mechanism has properties $\tilde{S}$ and O. Whether player 1's type is $+1$ or -1 does not affect the distribution of the message he receives (which has probability 0.5 of being L in both cases), and has no effect whatsoever on player 2's message. The corresponding correlated equilibrium distribution is not implementable by any mechanism that satisfies S and O, i.e., one in which both messages are unaffected by 1's type. The reason is that, in any correlated equilibrium with a mechanism that has that property, the expected payoff for the two types of player 1 must be equal. Otherwise, one of them could increase his payoff by mimicking the way the other type reacts to the message he receives.

[18] This verbal description of strategy correlated equilibrium is not entirely general, in that the recommendation that each player receives from the referee is an explicit strategy. In a more general setting, a device may send out messages that are arbitrary objects, and the translation into strategies for the receiving players is expressed by the correlated strategy.

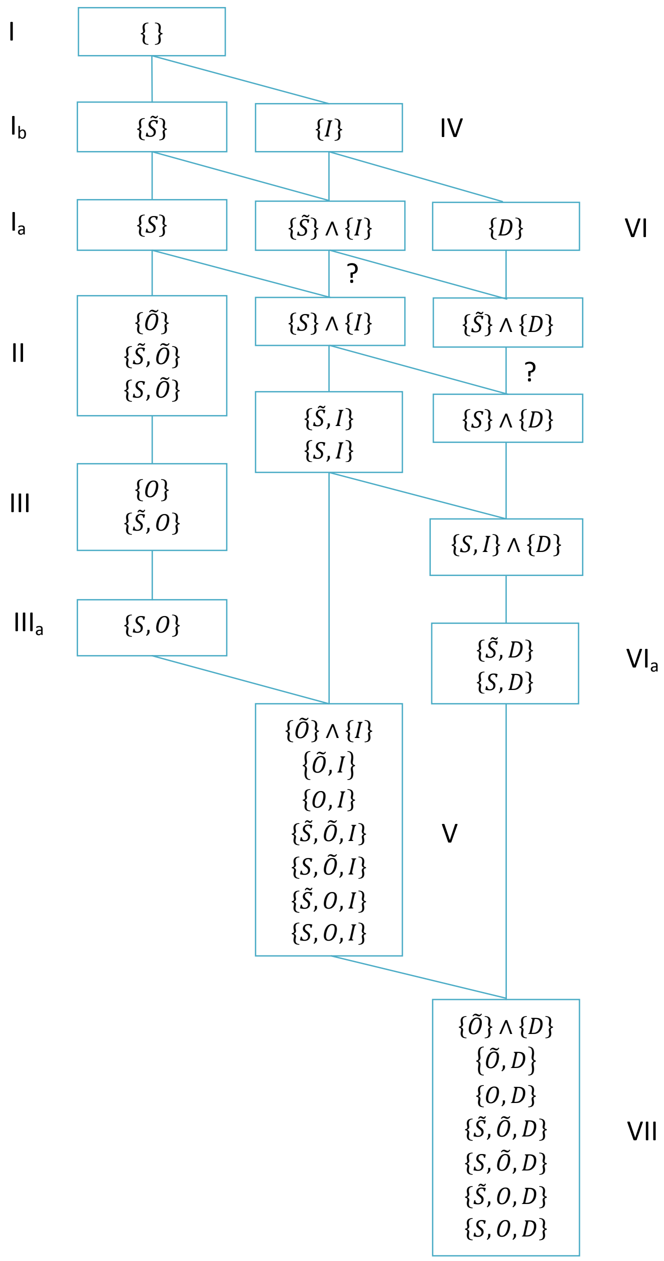

Figure 3. Hesse diagram of the different attributes of communication equilibrium distributions (MEDs) in Bayesian games, ordered by implication. Each attribute is represented by a box, and an implication relation is represented by a line. A conjunction symbol ∧ means that the MED is implementable both by a mechanism satisfying one property (or pair of properties) and by a mechanism satisfying the other property. A line marked ? represents an uncertain relation: it is not known whether the reverse implication also holds (in which case the two connected boxes should be coalesced).

3.3 Attributes of communication equilibrium distributions

A communication equilibrium in a Bayesian game is also a correlated equilibrium but the converse is generally not true. The incentive compatibility requirement for correlated equilibrium is that taking a different action than that prescribed by the correlated strategy cannot make a player better off. Communication equilibrium adds the requirement that reporting the types truthfully is incentive compatible. Obviously, the latter requirement has no bite if the mechanism ignores the players' type reports, i.e., if it has properties S and O. Thus, if a CED is implementable by a mechanism with these properties, it is automatically a MED. Somewhat less trivially, it is shown below that S, O-implementability is the *weakest* attribute of a CED in Figure 2 that guarantees that it is also a MED. Thus, for any list of properties of mechanisms that does not include S or O, some of the CEDs implementable by a mechanism with these properties are not MEDs. However, other such CEDs *are* MEDs, which raises the question of whether, also as MEDs, they are implementable by a mechanism with the same properties, i.e., whether there exists a communication equilibrium with such a mechanism that gives the distribution. Examples 7 and 8 below show that the answer to this question can be negative.

Implementability by a mechanism with a particular set of properties is an attribute of MEDs, just as for CEDs and CSDs. As in Sections 3.1 and 3.2, a basic question, for each such attribute of MEDs or a conjunction of several attributes, is which of the other attributes are implied by it. The answer is given by the Hesse diagram in Figure 3, which presents the implication relations among the various attributes of communication equilibrium distributions. Comparison with the diagram for correlated equilibrium distributions (Figure 2) shows that, among the attributes of distributions that are defined by a single set of properties of mechanisms, the implications relations for MEDs and CEDs are identical. However, this is not so for attributes that are defined by conjunction, with the result that some such attributes are equivalent in the case of CEDs but distinct for MEDs. For example, for MEDs, unlike for CEDs, the conjunction of S- and D-implementability does not imply S,I-implementability.

Example 7. *A communication equilibrium distribution that is S- as well as D-implementable but not S,I-implementable.* In a three-player Bayesian game, player 1 has two types, t_1' and t_1''. Players 2 and 3 both have the same two types, t' and t''. All type profiles except (t_1', t', t') may occur, and they have the same probability (1/7). Each player can play L or R. Player 1's payoff depends only on the type profile t and on the other players' actions. Specifically, it can be different from 0 only if $t = (t_1'', t', t')$, and in that case the payoff is given by the following symmetric matrix, in which the rows and columns correspond to the actions of players 2 and 3:

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} L \\ R \end{array} & \begin{pmatrix} 0 & -1 \\ -1 & 6 \end{pmatrix} \end{array}.$$

For player 2 and for player 3 the payoff is the sum of two numbers. The first number is 4 if player 1 plays L and players 2 and 3 have identical types; it is 4 also if player 1 plays R and players 2 and 3 have different types; and it is 0 otherwise. The second number is $1/2$ if the player's own action is R; and if the action is L, it is given by the following table, in which the rows describe the player's own type and the columns describe the types of the other two players:

	t_1', t'	t_1'', t'	t_1', t''	t_1'', t''
t'	N/A	1	0	0
t''	3	0	0	0

Consider the function that, for each type profile t, assigns to player 1 the action L or R if the types of players 2 and 3 are identical or different, respectively, and assigns to players 2 and 3 the actions specified by the following table, and which the rows and columns correspond to the player's own type and to that of the other player, respectively:

$$\begin{array}{c} \\ t' \\ t'' \end{array} \begin{array}{cc} t' & t'' \\ \hline \multicolumn{1}{|c}{L} & \multicolumn{1}{c|}{R} \\ \multicolumn{1}{|c}{L} & \multicolumn{1}{c|}{L} \\ \hline \end{array}.$$

With the mechanism with property D that, for each type profile t, tells each player the action that the above function assigns him, the correlated strategy of acting accordingly is a communication equilibrium. Player 1 cannot increase his payoff of 0, since there is no way he can make players 2 and 3 play R when they both have type t'. And for these players, a truthful type report is incentive compatible, since if (only) one of them lies, both players 2 and 3 lose the 4 they would get from a match between their types (identical or different) and player 1's action (L or R, respectively). In addition, for players 2 and 3, acting according to the coordinated strategy is incentive compatible. For a player of type t', doing so always guarantees him maximum payoff, and for type t'', deviating from the assigned action L to R would decrease the expected payoff by $(1/4 \times 3 - 1/2 =)\ 1/4$.

The communication equilibrium distribution described above is also implementable by a mechanism with property S. That mechanism sends to player 1 the same messages as above, and sends to each of the other two players i $(= 2{,}3)$ a message that depends on the others' types according to the table

$$\begin{array}{cccc} t_1', t' & t_1'', t' & t_1', t'' & t_1'', t'' \\ \hline \boldsymbol{Z}_i & L & R & R \end{array},$$

where $(\boldsymbol{Z}_2, \boldsymbol{Z}_3)$ is a pair of dependant random variables that equals (L, R) with probability 0.5 and (R, L) with probability 0.5. A communication equilibrium with this mechanism that gives the same distribution as the previous one is for each player to play according to the message he receives, unless he is of type t'', in which case he should play L. For a player of that type (t''), playing R would not increase the conditional expectation of the payoff, regardless of the message he receives. This is because, given that the received message is L or R, the conditional probability that the other players have types t_1' and t' is $1/3$ or $1/5$, respectively. Since both $1/3 \times 3$ and $1/5 \times 3$ are greater than $1/2$, deviation to R is unwarranted in both cases. The incentive compatibility of truthful type reports is proved by arguments similar to those used for the previous equilibrium.

There is no communication equilibrium with a mechanism with properties S and I that gives the above distribution. To see this, suppose that such a communication equilibrium exists. Since property S implies $\tilde{S}$, the distribution of the messages that player 3 gets from the mechanism only depends on the other players' types, so that it can be described by the table

$$\begin{array}{cccc} t_1', t' & t_1'', t' & t_1', t'' & t_1'', t'' \\ \hline P^1 & P^2 & P^3 & P^4 \end{array},$$

where P^1, P^2, P^3, P^4 are four probability measures on player 3's message space M_3. If the type of player 3 is t', he should play L or R if he receives any message in $\text{supp}(P^2)$ or in $\text{supp}(P^3) \cup \text{supp}(P^4)$, respectively. Therefore, these two subsets of M_3 must be disjoint. If the type of player 3 is t'', he should play L regardless of the message m_3 he receives. Deviation to R should not increase the conditional expectation of the player's payoff, which means that

$$\left(\frac{1}{2}-3\right)P^1(\{m_3\})+\frac{1}{2}P^2(\{m_3\})+\frac{1}{2}P^3(\{m_3\})+\frac{1}{2}P^4(\{m_3\})\leq 0.$$

Summing over all $m_3 \in \text{supp}(P^3) \cup \text{supp}(P^4)$ gives

$$-\frac{5}{2}P^1(\text{supp}(P^3)\cup\text{supp}(P^4))+0+\frac{1}{2}+\frac{1}{2}\leq 0.$$

It follows that if the type profile is (t_1', t', t'), the probability that player 3 plays R is at least $2/5$. The same is true for player 2. Therefore, by the assumed independence of the messages (property I), the probability that both 2 and 3 play R when the type profile is (t_1', t', t') is at least $4/25$. Since $4/25 \times 6 + 21/25 \times (-1) > 0$, this shows that player 1 has an incentive to misreport his type as t_1' when it is really t_1'', which contradicts the equilibrium assumption.

The MED in Examples 7 is not S, I-implementable, although it is (S- and D-implementable, and hence) S, I-implementable as a CED. The next example demonstrates another difference of this kind between MEDs and CEDs. The special significance of this example is that it shows that these solution concepts are not connected by a relation similar to that in Theorem 1, which concerns CEDs and CSDs. The example is taken from Gerardy (2004, Example 2).

Example 8. *A communication equilibrium distribution that is not D-implementable but does have that attribute as a correlated equilibrium distribution.* In a three-player Bayesian game, player 1 has two types, t_1' and t_1'', player 2 has two types, t_2' and t_2'', and player 3 has a single type. Types t_1'' and t_2'' cannot occur together, but all the other three type profiles are possible and equally probable. Players 1 and 2 have a single action, and player 3 has four actions: $a_3^1, a_3^2, a_3^3, a_3^4$. The four payoff vectors that correspond to the four actions, for each type profile, are given by the following table.

	t_2'	t_2''
t_1'	$(-1,-1,1),(1,1,0),(0,0,-1),(0,0,-1)$	$(1,0,0),(1,1,1),(-3,0,1),(1,0,-1)$
t_1''	$(0,1,0),(1,1,1),(0,1,-1),(0,-3,1)$	N/A

.

For each of the three possible type profiles there are either one or two actions for player 3 that yield him his maximum payoff of 1. There are four different ways to choose one such action for each type profile, and each such choice of actions specifies an S, D- (but not $\tilde{O}$-) implementable CED. However, it can be shown that only one of these four CEDs is also a MED, namely, the one in which player 3 chooses his first action a_3^1 if the other players' types are t_1' and t_2' and chooses the second action a_3^2 otherwise. In addition, in any communication equilibrium, player 3 randomizes exactly fifty-fifty between a_3^3 and a_3^4 if the impossible type combination (t_1'', t_2'') is reported; otherwise, truthful type reports are not incentive compatible. The above MED is therefore S, I- but not D-implementable.

The existence in Example 8 of three S, D-implementable CEDs that are not MEDs shows that this attribute of CEDs is insufficient to make them MEDs. Example 6 proves that $\tilde{S}, O$-implementability is also insufficient. This is because, in the CED considered in that example, the otherwise identical two types of player 1 receive different payoffs, which is impossible in a MED. Together, these examples prove that the only attributes of CEDs in Figure 2 that *necessarily* make them MEDs are those that imply S, O-implementability.

Example 8 also shows that a result similar to Corollary 1 does not hold for communication equilibrium distributions: the canonical strategy, with the canonical mechanism, is not necessarily a communication equilibrium. The reason for this fundamental difference between communication and correlated equilibria is that, in the former unlike the latter, the messages that the mechanism sends when it receives type reports that are patently not all truthful (since the profile lies outside the support of the distribution of type profiles) cannot be chosen arbitrarily. The mechanism's reaction to such reports has to induce actions that punish the player who lied about his type, whose identity may or may not be inferable. The feasibility of such a reaction may depend on the properties of the implementing mechanism. Hence the difference between Figure 2 and Figure 3.

The differences between the attributes of correlated equilibrium distributions and those of communication equilibrium distributions are not limited to differences between the respective Hesse diagrams. As Examples 7 and 8 show, they involve not only the number of attributes and the relations between them but also the placement of individual distributions. The same distribution may be assigned to either of two non-corresponding classes depending on whether it is viewed as a CED or as a MED. The classifications of CSDs, CEDs and MEDs, which are derived from the respective Hesse diagrams, are described in detail in the next subsection.

3.4 Classifications

A significant property of the collections of attributes of distributions that are described in the preceding three subsections is that each of them is closed under conjunctions. That is, each collection includes every attribute that can be defined as the conjunction of several of its elements, i.e., as the quality of possessing all of these attributes. This result is formally expressed by the following theorem, and it is proved by Lemmas 5, 6 and 7 in Sections 4, 5 and 6 below.

Theorem 2. The conjunction of any number of the attributes of CSDs in Figure 1, CEDs in Figure 2, or MEDs in Figure 3 is equivalent to one of the attributes in the same figure.

Each of the collections of attributes of CSDs, CEDs, and MEDs in the above figures is also a *lattice* with respect to the implication relation. That is, in each of the three Hesse diagrams, every two attributes have a greatest lower bound (or infimum) and a least upper bound (or supremum). The greatest lower bound, also called the *meet* of the two attributes, is the unique attribute in the diagram that (i) implies both attributes and (ii) is implied by every other attribute in the diagram that implies them. The least upper bound, also called the *join* of the two attributes, is the unique attribute that (i) is implied by each of the two attributes and (ii) implies every other attribute in the diagram that is implied by each of them. The meet and join operations are customarily denoted by ∧ and ∨ respectively. Theorem 2 implies that this notation is consistent with the use of ∧ in Figure 2 and Figure 3 as the symbol for logical conjunction. By the theorem, in each Hesse diagram, the conjunction of any two attributes is equivalent to some attribute in the same diagram. That attribute is clearly the meet of the first two: it implies each of them, and it is implied by any other attribute that does the same.

A CSD, CED, or MED in a Bayesian game usually has more than one of the attributes of distributions of that kind that are described in this paper. Specifically, if it has a particular attribute, then it also possesses every other attribute that is implied by the first one. For example, every random-profile distribution (attribute III in Figure 1) also has the conditional independence property (property II). However, as the following theorem shows, among all the attributes in Figure 1, Figure 2 or Figure 3 that a given CSD, CED or MED has, there is always one that implies all the others; it is its *strongest* attribute. Clearly, specifying the strongest attribute is equivalent to specifying the whole collection of attributes that the distribution possesses.

Theorem 3. For every correlated strategy distribution η, the collection of all the attributes in Figure 1 that η possesses includes one attribute that implies all the others. The same is true for correlated equilibrium distributions and for communication equilibrium distributions, except that for these the relevant attributes are those in Figure 2 and Figure 3, respectively.

Theorem 3 is an immediate corollary of the closedness under conjunctions. By Theorem 2, the conjunction of all the attributes that a distribution η possesses is equivalent to one of the attributes in the relevant Hesse diagram. Since equivalence means two-way implication, it follows that (i) η has that attribute, and (ii) the attribute implies all the other attributes that η possesses. Parenthetically, Theorem 3 does not simply follow from the observation that each of the three Hesse diagrams is a lattice (or vice versa). Removing VI_b, for example, from Figure 2 would invalidate the theorem, but the Hesse diagram would still represent a lattice.

As an illustration of Theorem 3, consider a CED that is implementable both by a mechanism with property $\tilde{O}$ and by a mechanism with property D. That CED must be a pure-equilibrium distribution. This is because the only attribute in Figure 2 that is stronger than (i.e., implies both) $\tilde{O}$-implementability and D-implementability is S, O, D-implementability.

Classification according to the strongest attribute partitions the collection of all CSDs into 7 nonempty and mutually disjoint classes. The partition for CEDs, which is finer than (that inherited from) the former, has 14 or 15 elements, and for MEDs the number of classes is 15, 16 or 17.[19] Each of these classes can be designated by the same roman number and (if applicable) subscript letter that designate the corresponding attribute in Figure 1, Figure 2 or Figure 3. For example, class II of CSDs consists of all the correlated strategy distributions with the conditional independence property which are *not* random-profile distributions.

3.5 Payoffs

The purpose of Example 6 is to demonstrate that the joint distributions of type and action profiles achievable by strategy correlated equilibria are not identical to those achievable by type correlated equilibria. It demonstrates that by showing that the two kinds of equilibria

[19] Although the ranges of possible cardinalities overlap, it follows from Example 7 and Lemma 22 below that the number of classes of MEDs is strictly greater than for CEDs.

may give different expected payoffs to certain types of players. Forges' (1993) demonstration of the nonequivalence of strategy correlated and type correlated equilibria, which is considerably more involved than Example 6, seemingly goes further by showing that even the *players'* payoffs, which combine those of all their types, may be different for the two kinds of equilibria. However, it follows from the next theorem that to study the effect of the properties of the implementing mechanism on the *correlated equilibrium payoffs* (CEPs), which are the n-tuples giving the players' expected payoffs in the correlated equilibria in an n-player Bayesian game, it is in fact not necessary to actually examine these payoffs, as Forges (1993, 2006) did. It suffices to solve the more tractable problem of CED implementability (Section 3.2). This is because any two kinds of mechanisms (of those considered in Figure 2) that do not implement the same CEDs *necessarily* also do not implement the same CEPs (and, obviously, vice versa). A similar relation exists between correlated strategy distributions (Section 3.1) and *correlated strategy payoffs* (CSPs), and between communication equilibrium distributions (Section 3.13.3) and *communication equilibrium payoffs* (MEPs). The proof of the theorem, which is given in Section 7, is constructive. It thus provides a means of automatically turning an example such as Example 6 into one that involves different payoff vectors (rather than just different joint distributions of types and actions).

Theorem 4. For any two subsets $\mathcal{P}$ and $\mathcal{Q}$ of the properties (of mechanisms) $\{S, \tilde{S}, O, \tilde{O}, D, I\}$, the proposition

$$\mathcal{P}\text{-implementability implies } \mathcal{Q}\text{-implementability}$$

holds for correlated strategy distributions, correlated equilibrium distributions or communication equilibrium distributions if and only if it holds for correlated strategy payoffs, correlated equilibrium payoffs or communication equilibrium payoffs, respectively. Moreover, the same is true with the premise "$\mathcal{P}$-implementability" replaced by "$\mathcal{P}'$-implementability and $\mathcal{P}''$-implementability and …", for any list $\mathcal{P}', \mathcal{P}''$, … of subsets of $\{S, \tilde{S}, O, \tilde{O}, D, I\}$.

The gist of Theorem 4 is that the Hesse diagrams in Figure 1, Figure 2 and Figure 3 apply not only to CSDs, CEDs and MEDs, respectively, but also to CSPs, CEPs and MEPs. Moreover, the classifications of CSPs, CEPs and MEDs by the properties of the implementing mechanisms are identical to the classifications of CSDs, CEDs and MEDs described in Section 3.4. Therefore, the notation used for the various classes of distributions may also be used for the corresponding classes of payoff vectors. For example, Class I of CEPs consists of all the correlated equilibrium payoffs (in specified Bayesian games) that are *not* $\tilde{S}$- or I-implementable.

This concludes the summary of the main results in this paper. The following sections present these results in detail and give their proofs.

4 Correlated Strategy Distributions

One of the goals of this paper is to identify new attributes of correlated strategy distributions. It would seem natural to base the classification on the intrinsic properties of

CSDs, that is, properties that can easily be expressed in terms of the distributions themselves, as in Lemma 2. Indeed, this is the prevalent approach in the literature. However, an alternative approach turns out to be quite fruitful. This approach, which forms the basis for this work, is to classify CSDs according to the properties of the mechanisms implementing them. Characterization in terms of intrinsic properties is the second step.

The classification of CSDs is based on the six properties of mechanisms described in Section 2.2, namely, S, $\tilde{S}$, O, $\tilde{O}$, D and I. Each subset $\mathcal{P}$ of these properties defines an attribute of CSDs, namely, $\mathcal{P}$-implementability. A CSD is $\mathcal{P}$-implementable if it is implementable by some mechanism with (all) the properties in $\mathcal{P}$. If $\mathcal{P}$ and $\mathcal{Q}$ are two subsets of properties, $\mathcal{P}$-implementability *implies* $\mathcal{Q}$-implementability if in every Bayesian game every $\mathcal{P}$-implementable CSD is also $\mathcal{Q}$-implementable. Shorthand for this relation is

$$\mathcal{P} \Rightarrow \mathcal{Q}. \tag{18}$$

A trivial sufficient condition for it is reverse inclusion, $\mathcal{P} \supseteq \mathcal{Q}$. $\mathcal{P}$-implementability and $\mathcal{Q}$-implementability are *comparable* if (18) or the reverse implication holds, and *equivalent* if both implications hold. Shorthand for equivalence is

$$\mathcal{P} \Leftrightarrow \mathcal{Q}$$

The connection between properties of mechanisms and attributes of CSDs can be extended by considering pairs of subsets of $\{S, \tilde{S}, O, \tilde{O}, D, I\}$. Each such pair, $\mathcal{P}$ and $\mathcal{P}'$, defines an attribute of CSDs , namely, the *conjunction* of $\mathcal{P}$-implementability and $\mathcal{P}'$-implementability, which is denoted by

$$\mathcal{P} \wedge \mathcal{P}'.$$

A CSD with this attribute is implementable both by a mechanism with the properties in $\mathcal{P}$ and by a (generally different) mechanism with the properties in $\mathcal{P}'$. However, Lemma 5 at the end of this section shows that no new attributes are defined this way, since every CSD as above is also implementable by a mechanism that has *both* the properties in $\mathcal{P}$ and those in $\mathcal{P}'$.

A simple and natural mechanism that implements any given CSD is its canonical mechanism (see Section 2.3). The following useful lemma identifies several attributes of CSDs that only depend on properties of their canonical mechanism.

Lemma 4. A CSD is $\tilde{O}$-, D- or I- implementable if and only if its canonical mechanism has property $\tilde{O}$, D or I, respectively.

Proof. Consider a CSD η and its canonical mechanism $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$. Let σ be any correlated strategy with a mechanism $\boldsymbol{m}'$ such that η is equal to the joint distribution of the random type profile $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ that is defined by a similar equation to (7) except that $\boldsymbol{m}'$ replaces $\boldsymbol{m}$. By definition of the canonical mechanism,

$$\boldsymbol{m}(t) \overset{d}{=} \left(\sigma_j(t_j, \boldsymbol{m}_j'(t))\right)_{j \in N}, \qquad t \in \operatorname{supp}(\eta_T). \tag{19}$$

If $\boldsymbol{m}'$ satisfies I, then the entries on the right-hand side of the equality are independent, and therefore (4) holds. A special property of the canonical mechanism is that (4) also holds for all $t \notin \mathrm{supp}(\eta_T)$, which proves that $\boldsymbol{m}$ satisfies I.

Another special property of the canonical mechanism is that, for every t and i, there is some t' with $(t_i, t'_{-i}) \in \mathrm{supp}(\eta_T)$ such that (2) holds, and hence (by (19))

$$\boldsymbol{m}_i(t) \overset{d}{=} \sigma_i(t_i, \boldsymbol{m}'_i(t_i, t'_{-i})).$$

If $\boldsymbol{m}'$ satisfies D, then the expression on the right-hand side has a degenerate distribution, which proves that $\boldsymbol{m}$ satisfies D. If $\boldsymbol{m}'$ satisfies $\tilde{O}$, then the distribution of the expression on the right-hand side is unaffected by replacing t' with an arbitrary type profile t^*, and the equation that results from this replacement proves that $\boldsymbol{m}$ also satisfies $\tilde{O}$. ■

Lemma 4 cannot unfortunately be extended to all attributes of CSDs. In particular, as indicated in Section 2.2.1, the canonical mechanism of an O-implementable CSD does not necessarily have property O. However, every O-implementable CSD η is implementable by a mechanism $\widetilde{\boldsymbol{m}} = (\widetilde{\boldsymbol{m}}(t))_{t\in T}$ with property O that is *functionally indistinguishable* from the canonical mechanism $\boldsymbol{m} = (\boldsymbol{m}(t))_{t\in T}$, in that

$$\widetilde{\boldsymbol{m}}(t) \overset{d}{=} \boldsymbol{m}(t), \qquad t \in \mathrm{supp}(\eta_T). \tag{20}$$

Such a mechanism can be constructed by taking any mechanism $\boldsymbol{m}' = (\boldsymbol{m}'(t))_{t\in T}$ with property O that implements η and a corresponding correlated strategy σ, and defining

$$\widetilde{\boldsymbol{m}}_i(t) = \sigma_i(t_i, \boldsymbol{m}'_i(t)), \qquad i \in N, t \in T.$$

It follows from (19) that this mechanism is functionally indistinguishable from the canonical one, which implies that η is obtained also from using the canonical strategy with the mechanism $\widetilde{\boldsymbol{m}}$.

4.1 Intrinsic characterizations

This subsection presents intrinsic characterizations for several attributes of CSDs, which represent an alternative to definitions by properties of the implementing mechanisms. The first two propositions are reworded versions of results presented in Section 3.1.1.

Proposition 1. A CSD η is S, O, I-implementable if and only if the following condition holds for some (equivalently, every) pair of random variables $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \dots, \boldsymbol{t}_n)$ and $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ whose joint distribution is η:

(i) For each player i, $\boldsymbol{a}_i$ and $(\boldsymbol{t}_{-i}, \boldsymbol{a}_{-i})$ are conditionally independent, given $\boldsymbol{t}_i$.

A CSD is S, O, D-implementable if and only if it satisfies the stronger condition in which (i) is replaced by:

(ii) For each player i, the conditional distribution of $\boldsymbol{a}_i$, given $\boldsymbol{t}_i$, is degenerate.

Proof. To prove the sufficiency of the condition for S, O, I-implementability, let $\boldsymbol{m}$ be the canonical mechanism of η and $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$ the canonical random action profile. Suppose that

(i) holds. For every type profile $t \in \mathrm{supp}(\eta_T)$ and action profile a,

$$\begin{aligned}\Pr(\boldsymbol{a} = a|\boldsymbol{t} = t) &= \Pr(\boldsymbol{a}_1 = a_1|\boldsymbol{a}_{-1} = a_{-1}, \boldsymbol{t} = t)\Pr(\boldsymbol{a}_{-1} = a_{-1}|\boldsymbol{t} = t) \\ &= \Pr(\boldsymbol{a}_1 = a_1|\boldsymbol{t} = t)\Pr(\boldsymbol{a}_{-1} = a_{-1}|\boldsymbol{t} = t) = \cdots \\ &= \Pr(\boldsymbol{a}_1 = a_1|\boldsymbol{t} = t)\Pr(\boldsymbol{a}_2 = a_2|\boldsymbol{t} = t)\cdots\Pr(\boldsymbol{a}_n = a_n|\boldsymbol{t} = t),\end{aligned} \tag{21}$$

where the second equality follows from (i) and the subsequent equalities follow from using an identical trick for the other entries of a. This proves that $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ are conditionally independent, given $\boldsymbol{t}$. By (i), for every type profile t' with $(t_i, t'_{-i}) \in \mathrm{supp}(\eta_T)$, each expression in (21) of the form $\Pr(\boldsymbol{a}_i = a_i|\boldsymbol{t} = t)$ is equal to $\Pr(\boldsymbol{a}_i = a_i|\boldsymbol{t} = (t_i, t'_{-i}))$. Therefore, since $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$, (21) and the second part of property (5) of the canonical mechanism imply that, for $t \in \mathrm{supp}(\eta_T)$, $a \in A$ and *any* $t' \in T$,

$$\Pr(\boldsymbol{a} = a|\boldsymbol{t} = t) = \Pr(\boldsymbol{m}(t) = a) = \prod_{i=1}^{n} \Pr(\boldsymbol{m}_i(t_i, t'_{-i}) = a_i). \tag{22}$$

Assume, without loss of generality, that the indexing of the players' types (see Section 2.1) is such that $t^1 = (t_1^1, t_2^1, \dots, t_n^1) \in \mathrm{supp}(\eta_T)$. For each t, let $\overline{\boldsymbol{m}}(t) = (\overline{\boldsymbol{m}}_1(t), \overline{\boldsymbol{m}}_2(t), \dots \overline{\boldsymbol{m}}_n(t))$ be the random variable with values in $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ that is defined by

$$\overline{\boldsymbol{m}}_i(t) = \left(\boldsymbol{m}_i(t_i^1, t_{-i}^1), \boldsymbol{m}_i(t_i^2, t_{-i}^1), \dots\right), \qquad i \in N. \tag{23}$$

The message that each player i receives from the mechanism $\overline{\boldsymbol{m}} = (\overline{\boldsymbol{m}}(t))_{t \in T}$ is a pure strategy of the form $(a_i^1, a_i^2, \dots) \in A_i^{T_i}$, where each action a_i^k coincides with the message that the player would receive from the canonical mechanism $\boldsymbol{m}$ if his type were t_i^k (the player's kth type) and the other players' types were given by t_{-i}^1. Since this description does not involve in any way the actual type profile t, the mechanism $\overline{\boldsymbol{m}}$ satisfies S and O. To prove that it also satisfies I, it has to be shown that for any $a_1^1, a_1^2, \dots \in A_1$, $a_2^1, a_2^2, \dots \in A_2$, …, $a_n^1, a_n^2, \dots \in A_n$,

$$\begin{aligned}&\Pr(\boldsymbol{m}_1(t_1^1, t_{-1}^1) = a_1^1, \boldsymbol{m}_1(t_1^2, t_{-1}^1) = a_1^2, \dots; \boldsymbol{m}_2(t_2^1, t_{-2}^1) = a_2^1, \boldsymbol{m}_2(t_2^2, t_{-2}^1) = a_2^2, \dots; \dots) \\ &\quad = \prod_{i=1}^{n} \Pr(\boldsymbol{m}_i(t_i^1, t_{-i}^1) = a_i^1, \boldsymbol{m}_i(t_i^2, t_{-i}^1) = a_i^2, \dots).\end{aligned} \tag{24}$$

By (6), the left-hand side is equal to

$$\Pr(\boldsymbol{m}(t^1) = (a_1^1, a_2^1, \dots, a_n^1)) \cdot \prod_{i=1}^{n} \left(\Pr\left(\boldsymbol{m}_i(t_i^2, t_{-i}^1) = a_i^2\right) \Pr\left(\boldsymbol{m}_i(t_i^3, t_{-i}^1) = a_i^3\right) \cdots\right).$$

By the second equality in (22) (used with $t = t' = t^1$), this proves that

$$\begin{aligned}&\Pr(\boldsymbol{m}_1(t_1^1, t_{-1}^1) = a_1^1, \boldsymbol{m}_1(t_1^2, t_{-1}^1) = a_1^2, \dots; \boldsymbol{m}_2(t_2^1, t_{-2}^1) = a_2^1, \boldsymbol{m}_2(t_2^2, t_{-2}^1) = a_2^2, \dots; \dots) \\ &\quad = \prod_{i=1}^{n} \prod_{k \ge 1} \Pr(\boldsymbol{m}_i(t_i^k, t_{-i}^1) = a_i^k).\end{aligned} \tag{25}$$

By (6) again, (25) proves (24), so that the mechanism $\overline{\boldsymbol{m}}$ satisfies I. To prove that it implements the CSD η, define a correlated strategy $\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2, \dots, \bar{\sigma}_n)$ with the mechanism

$\bar{\boldsymbol{m}}$ by

$$\bar{\sigma}_i(t_i^k, (a_i^1, a_i^2, \dots)) = a_i^k, \qquad i \in N, k = 1,2, \dots. \tag{26}$$

Thus, according to $\bar{\sigma}_i$, of all the entries in the message, player i takes the one corresponding to his actual type. It has to be shown that

$$(\boldsymbol{t}, \boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n) \overset{d}{=} (\boldsymbol{t}, \bar{\sigma}_1(\boldsymbol{t}_1, \bar{\boldsymbol{m}}_1(t)), \bar{\sigma}_2(\boldsymbol{t}_2, \bar{\boldsymbol{m}}_2(t)), \dots, \bar{\sigma}_n(\boldsymbol{t}_n, \bar{\boldsymbol{m}}_n(t))).$$

For this, it suffices to show that, for every $t \in \mathrm{supp}(\eta_T)$ and $a \in A$,

$$\begin{aligned}&\Pr(\boldsymbol{a} = a | \boldsymbol{t} = t)\\ &= \Pr(\bar{\sigma}_1(\boldsymbol{t}_1, \bar{\boldsymbol{m}}_1(\boldsymbol{t})) = a_1, \bar{\sigma}_2(\boldsymbol{t}_2, \bar{\boldsymbol{m}}_2(\boldsymbol{t})) = a_2, \dots, \bar{\sigma}_n(\boldsymbol{t}_n, \bar{\boldsymbol{m}}_n(\boldsymbol{t})) = a_n | \boldsymbol{t} = t).\end{aligned}$$

By (22), this equation holds if and only if

$$\begin{aligned}&\prod_{i=1}^{n} \Pr\big(\boldsymbol{m}_i(t_i, t_{-i}^1) = a_i\big)\\ &\qquad = \Pr(\bar{\sigma}_1(t_1, \bar{\boldsymbol{m}}_1(t)) = a_1, \bar{\sigma}_2(t_2, \bar{\boldsymbol{m}}_2(t)) = a_2, \dots, \bar{\sigma}_n(t_n, \bar{\boldsymbol{m}}_n(t)) = a_n).\end{aligned}$$

By (26) and (23), the left-hand side is equal to

$$\prod_{i=1}^{n} \Pr(\bar{\sigma}_i(t_i, \bar{\boldsymbol{m}}_i(t)) = a_i),$$

and by property I of the mechanism $\bar{\boldsymbol{m}}$, the right-hand side is also equal to this product. Therefore, the equality holds.

Establishing the sufficiency of the condition for S, O, D-implementability only requires the following short addition to the above proof. Suppose that (ii) (rather than the weaker condition (i)) holds. It has to be shown that the mechanism $\bar{\boldsymbol{m}}$ satisfies D (rather than only I). By (23), it suffices to show that for every $t' \in T$ and $i \in N$ the distribution of the random variable $\bar{\boldsymbol{m}}_i(t')$ is degenerate. By the second part of property (5) of the canonical mechanism, it suffices to restrict attention to type profiles t' in $\mathrm{supp}(\eta_T)$, for which the distribution is equal to the conditional distribution of $\boldsymbol{m}_i(\boldsymbol{t})$, given that $\boldsymbol{t} = t'$. Since $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$, by (ii) that distribution is degenerate.

It remains to prove the necessity of the conditions in the proposition. Every CED is the joint distribution of pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a}$ such that (7) holds for some correlated strategy σ with a mechanism $\boldsymbol{m}$. Moreover, $\boldsymbol{t}$ and $\boldsymbol{m}$ are independent, and therefore

$$\Pr(\boldsymbol{a}_i = a_i | \boldsymbol{t} = t) = \Pr(\sigma_i(t_i, \boldsymbol{m}_i(t)) = a_i), \qquad i \in N, t \in \mathrm{supp}(\eta_T), a \in A. \tag{27}$$

It follows from (27) that if the mechanism $\boldsymbol{m}$ satisfies V, then the probability on the left-hand side is either 0 or 1, and if $\boldsymbol{m}$ satisfies $\tilde{O}$, then, for every $i \in N$, $t \in \mathrm{supp}(\eta_T)$ and $a \in A$,

$$\Pr(\boldsymbol{a}_i = a_i | \boldsymbol{t} = t) = \Pr(\boldsymbol{a}_i = a_i | \boldsymbol{t}_i = t_i). \tag{28}$$

If $\boldsymbol{m}$ satisfies I, then (in view of (7)) the actions $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ are conditionally independent, given $\boldsymbol{t}$, and therefore, for every $i \in N$, $t \in \operatorname{supp}(\eta_T)$ and $a \in A$,

$$\Pr(\boldsymbol{a} = a | \boldsymbol{t} = t) = \Pr(\boldsymbol{a}_i = a_i | \boldsymbol{t} = t) \Pr(\boldsymbol{a}_{-i} = a_{-i} | \boldsymbol{t} = t).$$

Multiplying both sides by the conditional probability $\Pr(\boldsymbol{t}_{-i} = t_{-i} | \boldsymbol{t}_i = t_i)$ gives

$$\Pr(\boldsymbol{t}_{-i} = t_{-i}, \boldsymbol{a} = a | \boldsymbol{t}_i = t_i) = \Pr(\boldsymbol{a}_i = a_i | \boldsymbol{t} = t) \Pr(\boldsymbol{t}_{-i} = t_{-i}, \boldsymbol{a}_{-i} = a_{-i} | \boldsymbol{t}_i = t_i). \quad (29)$$

Therefore, if a CSD η is both $\tilde{O}$- and I-implementable (and, a fortiori, if it is S, O, I-implementable), then every $\boldsymbol{t}$ and $\boldsymbol{a}$ whose joint distribution is η satisfy (28) and (29), and hence also

$$\Pr(\boldsymbol{t}_{-i} = t_{-i}, \boldsymbol{a} = a | \boldsymbol{t}_i = t_i) = \Pr(\boldsymbol{a}_i = a_i | \boldsymbol{t}_i = t_i) \Pr(\boldsymbol{t}_{-i} = t_{-i}, \boldsymbol{a}_{-i} = a_{-i} | \boldsymbol{t}_i = t_i),$$

for every $i \in N, t \in \operatorname{supp}(\eta_T)$ and $a \in A$. This property is equivalent to condition (ii). If η is in addition D-implementable (and, a fortiori, if it is S, O, D-implementable), then the expression on the right-hand side of (28) is either 0 or 1, which gives (i). ∎

Proposition 2. A CSD η is S, O-implementable if and only if there is a probability measure μ on $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ such that

$$\eta(\{(t, a)\}) = \eta_T(\{t\}) \cdot \mu^t(\{a\}), \quad t \in T, a \in A, \quad (30)$$

where η_T is the distribution of type profiles and μ^t is the marginal measure defined by (14).

Proof. To prove the sufficiency of the condition, suppose that a measure μ as above exists for a CSD η. Restrict μ to its support R, and let the random variable $\boldsymbol{r}$ be the identity map on R. By construction, $\boldsymbol{r}$ is independent of the random type profile $\boldsymbol{t}$. Define a mechanism $\boldsymbol{m} = (\boldsymbol{m}(t))_{t \in T}$ by

$$\boldsymbol{m}_i(t) = \boldsymbol{r}, \qquad i \in N, t \in T. \quad (31)$$

This mechanism clearly has properties S and O. The message space of each player is R, each element r of which is a pure-strategy profile $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$ (where, for each i and k, a_i^k is the action prescribed to player i's kth type). Define a correlated strategy σ with the mechanism $\boldsymbol{m}$ by

$$\sigma_i(t_i^k, r) = a_i^k, \qquad i \in N, k = 1,2, \dots. \quad (32)$$

Thus, the strategy for each player is simply to take the action prescribed to his actual type. It has to be shown that the joint distribution of $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ corresponding to σ is η. By (30), this means that the following has to be established:

$$\Pr(\boldsymbol{a} = a | \boldsymbol{t} = t) = \mu^t(\{a\}), \qquad t \in \operatorname{supp}(\eta_T), a \in A. \quad (33)$$

By (7) and (31), for any type profile $t = (t_1^{j_1}, t_2^{j_2}, \dots, t_n^{j_n})$,

$$\Pr(\boldsymbol{a} = a | \boldsymbol{t} = t) = \Pr\left((\sigma_1(t_1^{j_1}, \boldsymbol{r}), \sigma_2(t_2^{j_2}, \boldsymbol{r}), \dots, \sigma_n(t_n^{j_n}, \boldsymbol{r})) = a\right).$$

By (32), the right-hand side is the μ-measure of the set of all pure-strategy profiles $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$ with $(a_1^{j_1}, a_2^{j_2}, \dots, a_n^{j_n}) = a$, which by definition (Eq. (14)) is equal to $\mu^t(\{a\})$. Thus, (33) holds, so that $\boldsymbol{m}$ implements the CSD η.

To prove the necessity of the condition for S, O-implementability, consider a CSD η that is equal to the joint distribution of a pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a}$ such that (7) holds for a correlated strategy σ with a mechanism $\boldsymbol{m}$ that satisfies O (and may or may not satisfy S). Fix a type profile t'. The random variable

$$\left(\left(\sigma_1(t_1, \boldsymbol{m}_1(t_1, t'_{-1}))\right)_{t_1 \in T_1}, \left(\sigma_2(t_2, \boldsymbol{m}_2(t_2, t'_{-2}))\right)_{t_2 \in T_2}, \dots, \left(\sigma_n(t_n, \boldsymbol{m}_n(t_n, t'_{-n}))\right)_{t_n \in T_n}\right)$$

returns values in $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$, i.e., pure-strategy profiles. Its distribution μ is given by

$$\begin{aligned}\mu(\{(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots)\}) = \Pr(&\sigma_1(t_1^1, \boldsymbol{m}_1(t_1^1, t'_{-1})) = a_1^1, \sigma_1(t_1^2, \boldsymbol{m}_1(t_1^2, t'_{-1})) \\ &= a_1^2, \dots; \sigma_2(t_2^1, \boldsymbol{m}_2(t_2^1, t'_{-2})) = a_2^1, \sigma_2(t_2^2, \boldsymbol{m}_2(t_2^2, t'_{-2})) = a_2^2, \dots; \dots).\end{aligned}$$

For every type profile $t = (t_1^{j_1}, t_2^{j_2}, \dots, t_n^{j_n})$ and action profile $a = (a_1, a_2, \dots a_n)$,

$$\begin{aligned}&\mu\left(\left\{(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots) \in A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n} \,\middle|\, (a_1^{j_1}, a_2^{j_2}, \dots, a_n^{j_n}) = a\right\}\right) \\ &= \Pr\left(\sigma_1(t_1^{j_1}, \boldsymbol{m}_1(t_1^{j_1}, t'_{-1})) = a_1, \sigma_2(t_2^{j_2}, \boldsymbol{m}_2(t_2^{j_2}, t'_{-2})) = a_2, \dots, \sigma_n(t_n^{j_n}, \boldsymbol{m}_n(t_n^{j_n}, t'_{-n}))\right) \\ &= \Pr(\sigma_1(\boldsymbol{t}_1, \boldsymbol{m}_1(\boldsymbol{t})) = a_1, \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(\boldsymbol{t})) = a_2, \dots, \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(\boldsymbol{t})) = a_n | \boldsymbol{t} = t),\end{aligned}$$

where the last equality uses the assumption that $\boldsymbol{m}$ has property O. By (14) and (7), this shows that (33) holds, which gives (30). ■

Proposition 3. A CSD η is $\tilde{O}$-implementable if and only if it has the conditional independence property.

Proof. In view of Lemma 4, it suffices to show that the canonical mechanism $\boldsymbol{m}$ of η has property $\tilde{O}$ if and only if the condition in Definition 3 holds for the random type profile $\boldsymbol{t}$ and the canonical random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$. The condition in the definition is the requirement that, for every player i and type t_i for that player,

$$\Pr\left(\boldsymbol{a}_i = a_i \middle| \boldsymbol{t} = (t_i, t'_{-i})\right) = \Pr\left(\boldsymbol{a}_i = a_i \middle| \boldsymbol{t} = (t_i, t''_{-i})\right), \qquad a_i \in A_i \tag{34}$$

for all type profiles t' and t'' with $(t_i, t'_{-i}), (t_i, t''_{-i}) \in \operatorname{supp}(\eta_T)$. Since $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$, (34) is equivalent to

$$\boldsymbol{m}_i(t_i, t'_{-i}) \overset{d}{=} \boldsymbol{m}_i(t_i, t''_{-i}). \tag{35}$$

It follows from the second part of property (5) of the canonical mechanism that (35) holds for all type profiles t' and t'' with $(t_i, t'_{-i}), (t_i, t''_{-i}) \in \operatorname{supp}(\eta_T)$ if and only if it holds for *all* t' and t''. This is so for every player i and type t_i if and only if $\boldsymbol{m}$ has property $\tilde{O}$. ■

Proposition 4. A CSD η is I-implementable if and only if the following condition holds for some (equivalently, every) pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ whose joint distribution is η:

(i) $\{\boldsymbol{a}_j\}_{j\in N}$ are conditionally independent, given $\boldsymbol{t}$.

A CSD is D-implementable if and only if it satisfies the stronger condition in which (i) is replaced by:

(ii) The conditional distribution of $\boldsymbol{a}$ given $\boldsymbol{t}$ is degenerate.

Proof. In view of Lemma 4, it suffices to show that the canonical mechanism $\boldsymbol{m}$ of η has property I or D if and only if (i) or (ii), respectively, holds for the random type profile $\boldsymbol{t}$ and the canonical random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$. In other words, condition (4) or (3) holds for all type profiles t if and only if the condition holds for all $t \in \operatorname{supp}(\eta_T)$. This equivalence is implied by property (5) of the canonical mechanism. ■

4.2 Equivalences

This subsection identifies equivalent formulations for the attributes of CSDs considered in the previous subsection.

Proposition 5. For CSDs, $\{S\} \Leftrightarrow \{\tilde{S}\} \Leftrightarrow \{\,\}$, $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$, $\{S, O\} \Leftrightarrow \{\tilde{S}, O\} \Leftrightarrow \{O\}$, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\} \Leftrightarrow \{I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\} \Leftrightarrow \{D\}$.

Proof. Since property S of mechanisms implies $\tilde{S}$, and therefore $\{S\} \Rightarrow \{\tilde{S}\} \Rightarrow \{\,\}$, to prove that these three attributes are equivalent it suffices to show that every CSD is S-implementable.

Let $\boldsymbol{m}$ be the canonical mechanism of a CSD η. It implements η but does not necessarily have property S (see Section 2.2.1). To obtain an implementing mechanism that has that property, let $\vec{\boldsymbol{m}}(t) = (\vec{\boldsymbol{m}}_1(t), \vec{\boldsymbol{m}}_2(t), \dots, \vec{\boldsymbol{m}}_n(t))$ for each type profile t be the random variable with values in $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ that is defined by

$$\vec{\boldsymbol{m}}_i(t) = \left(\boldsymbol{m}_i(t_i^1, t_{-i}), \boldsymbol{m}_i(t_i^2, t_{-i}), \dots\right), \qquad i \in N. \tag{36}$$

This definition differs from that in (23) in that the partial type profile on the right-hand side is t_{-i} rather than the constant one t_{-i}^1. Hence, the mechanism $\vec{\boldsymbol{m}} = (\vec{\boldsymbol{m}}(t))_{t\in T}$ only satisfies S. With this mechanism, the correlated strategy $\bar{\sigma}$ defined in (26) gives the canonical random action profile:

$$\bar{\sigma}_i(\boldsymbol{t}_i, \vec{\boldsymbol{m}}_i(\boldsymbol{t})) = \boldsymbol{m}_i(\boldsymbol{t}), \qquad i \in N. \tag{37}$$

This proves that the mechanism $\vec{\boldsymbol{m}}$ implements η.

The proofs that $\{S, I\} \Leftrightarrow \{\tilde{S}, I\} \Leftrightarrow \{I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\} \Leftrightarrow \{D\}$ are very similar, and only require the following additions to the above proof.

If the CSD η is I- or D-implementable, then by Lemma 4 the canonical mechanism $\boldsymbol{m}$ has property I or D, respectively. Since a canonical mechanism also satisfies (6), the random variables

$$\{\boldsymbol{m}_j(t)\}_{j\in N, t\in T} \tag{38}$$

are independent in both cases, and in the case of D-implementability their distributions are moreover degenerate. It follows that, for every type profile t, the same is true for the n random variables

$$\{\vec{\boldsymbol{m}}_j(t)\}_{j\in N},$$

each of which is a vector whose entries are a subset of the random variables in (38), such that these n subsets are disjoint. This shows that, if η is I- or D-implementable, then the mechanism $\vec{\boldsymbol{m}}$ satisfies I or D, respectively, as well as S.

To prove that $\{S,\tilde{O}\} \Leftrightarrow \{\tilde{S},\tilde{O}\} \Leftrightarrow \{\tilde{O}\}$, it suffices to show that $\tilde{O}$-implementability implies $S,\tilde{O}$-implementability. In fact, in view of Lemma 4, it suffices to show that if the canonical mechanism $\boldsymbol{m}$ satisfies $\tilde{O}$, then $\vec{\boldsymbol{m}}$ also satisfies $\tilde{O}$. If the former condition holds, then, for every player i and type profiles t' and t'', (35) holds for all types t_i, which by (6) implies

$$\left(\boldsymbol{m}_i(t_i^1,t'_{-i}),\boldsymbol{m}_i(t_i^2,t'_{-i}),\dots\right) \stackrel{D}{=} \left(\boldsymbol{m}_i(t_i^1,t''_{-i}),\boldsymbol{m}_i(t_i^2,t''_{-i}),\dots\right).$$

Thus, $\vec{\boldsymbol{m}}$ has property $\tilde{O}$.

To prove that $\{S,O\} \Leftrightarrow \{\tilde{S},O\} \Leftrightarrow \{O\}$, it suffices to show that $\{O\} \Rightarrow \{S,O\}$. This is shown in the proof of Proposition 2, where it is proved that the existence of a measure μ as in that proposition implies S,O-implementability and it is implied by O-implementability. ■

Proposition 6. For CSDs, $\{S,O,I\} \Leftrightarrow \{\tilde{S},O,I\} \Leftrightarrow \{S,\tilde{O},I\} \Leftrightarrow \{\tilde{S},\tilde{O},I\} \Leftrightarrow \{O,I\} \Leftrightarrow \{\tilde{O},I\} \Leftrightarrow (\{\tilde{O}\} \wedge \{I\})$ and $\{S,O,D\} \Leftrightarrow \{\tilde{S},O,D\} \Leftrightarrow \{S,\tilde{O},D\} \Leftrightarrow \{\tilde{S},\tilde{O},D\} \Leftrightarrow \{O,D\} \Leftrightarrow \{\tilde{O},D\} \Leftrightarrow (\{\tilde{O}\} \wedge \{D\})$.

Proof. It clearly suffices to show that $(\{\tilde{O}\} \wedge \{I\}) \Rightarrow \{S,O,I\}$ and $(\{\tilde{O}\} \wedge \{D\}) \Rightarrow \{S,O,D\}$. As shown in the last part of the proof of Proposition 1, every CSD that is both $\tilde{O}$- and I-implementable, or both $\tilde{O}$- and D-implementable, respectively, is the joint distribution of a pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a}$ that satisfies condition (i) or (ii) in Proposition 1. Therefore, by that proposition, in the first case the CSD is also S,O,I-implementable, and in the second case it is S,O,D-implementable. ■

4.3 Implications

Propositions 5 and 6 identify seven attributes of correlated strategy distributions that are defined by subsets of the six properties of mechanisms defined in Section 2.2. Figure 1 presents these classes as well as certain trivial implication relations among them, which all follow immediately from relations between properties of mechanisms. To prove that the figure presents a complete picture of the implication relations between attributes of CSDs, it remains to show that implications additional to those shown do not hold, so that, in particular, none of the seven attributes is equivalent to another. For this, the following four propositions are required.

Proposition 7. For CSDs, $\{S,O,I\} \nRightarrow \{D\}$.

Proof. It suffices to consider any complete information game, that is, a game where every player has only one type, with a mixed-strategy profile that is not pure. ■

Proposition 8. For CSDs, $\{S, D\} \nRightarrow \{\tilde{O}\}$.

Proof. In a two-player Bayesian game in which player 1 has a single type and two actions and player 2 has a single action and two types, consider a correlated strategy distribution in which player 1 takes his first or second action if player 2 is of the first or second type, respectively. This CSD is implementable by mechanism that simply tells player 1 the type of player 2, and thus satisfies S and D. However, the CSD is not $\tilde{O}$-implementable, since with a mechanism that satisfies $\tilde{O}$, player 1 cannot possibly know player 2's type. ■

Proposition 9. For CSDs, $\{\tilde{O}\} \nRightarrow \{O\}$.

Proof. By Example 1, there exists a CSD that has the conditional independence property but is not S, O-implementable. By Proposition 3, that CSD is $\tilde{O}$-implementable, and by Proposition 5, it is not O-implementable. ■

Proposition 10. For CSDs, $\{S, O\} \nRightarrow \{I\}$.

Proof. In a complete information game, properties S and O automatically hold for every mechanism, but a CSD is I-implementable only if the players' actions are independent. ■

Proposition 7 proves that attribute V only implies the other attributes in Figure 1 that the diagram indicates it implies (in other words, it does not imply VI or VII). Proposition 8 proves the same for attribute VI. These two results prove that attribute IV (which is implied by both V and VI) only implies attribute I, and therefore the latter does not imply II. Proposition 9 proves that II does not imply III. Proposition 10 proves that attribute III only implies the (two) attributes that the diagram indicates it implies, which establishes the same for attribute II and for attribute I.

Since, for mechanisms, property S implies $\tilde{S}$, property O implies $\tilde{O}$, and D implies I, there are only 27 relevant subsets of $\{S, \tilde{S}, O, \tilde{O}, D, I\}$, which all appear in Figure 1. Therefore, there are no additional attributes of CSDs that can be described by single subsets of the six properties of mechanisms. The following lemma shows that the same is true for pairs (hence also triplets, etc.) of sets of properties of mechanisms: no additional attributes of CSDs can be defined by them.

Lemma 5. For CSDs, for every two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$(\mathcal{P} \wedge \mathcal{Q}) \Leftrightarrow \mathcal{P} \cup \mathcal{Q}. \tag{39}$$

Proof (an outline). Proposition 6 proves the special case of (39) in which $\mathcal{P} = \{\tilde{O}\}$ and $\mathcal{Q}$ is either $\{D\}$ or $\{I\}$. By inspection of Figure 1, every other case follows from one of these two. ■

As indicated in Section 3.2.2 (see also Section 5.3), for correlated equilibrium distributions a similar result to Lemma 5 does not hold. In other words, the requirement of incentive compatibility may invalidate the equivalence (39).

5 Correlated Equilibrium Distributions

The analysis of correlated strategy distributions in the previous section is a first step in the analysis of correlated equilibrium distributions. The former concerns qualitative differences between distributions that reflect the limited capabilities of the implementing mechanisms. The latter also incorporates the constraints inherent in the incentive compatibility requirement. Whereas in the case of CSDs the limiting factor is the mechanism's ability to transmit information to players, in the case of CEDs its ability to do so selectively also comes into play.

As for CSDs, each subset $\mathcal{P}$ of the six properties of mechanisms described in Section 2.2 defines an attribute of correlated strategy distributions, namely, $\mathcal{P}$-implementability. A CED has this attribute if it is implementable by some mechanism with (all) the properties in $\mathcal{P}$. Note that in the present context implementability has a different meaning than for CSDs (Section 4). Here, the correlated strategy involved is required to be a correlated equilibrium. Thus, an expression like (18) has a different meaning for CSDs and CEDs. Wherever confusion is possible, the generic implication sign may be replaced by the more explicit one $\underset{\text{CSD}}{\Rightarrow}$ or $\underset{\text{CED}}{\Rightarrow}$. The following proposition shows that the second of these relations is in a sense stronger than the first one.

Proposition 11. For every two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q} \text{ implies } \mathcal{P} \underset{\text{CSD}}{\Rightarrow} \mathcal{Q}. \tag{40}$$

Moreover, the same is true with $\mathcal{P}$ replaced by by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$,[20] for any list $\mathcal{P}', \mathcal{P}''$, … of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$.

Proof. It has to be shown that (i) $\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}$ and (ii) $\mathcal{P} \underset{\text{CSD}}{\not\Rightarrow} \mathcal{Q}$ are contradictory. Condition (i) means that, in every Bayesian game, every $\mathcal{P}$-implementable CED is also $\mathcal{Q}$-implementable. Condition (ii) means that there is some CSD in some Bayesian game that is $\mathcal{P}$- but not $\mathcal{Q}$-implementable. Without loss of generality, the payoff functions in that game (which are irrelevant for CSD implementability) are identically zero. However, this means that every correlated strategy in the game is a correlated equilibrium and vice versa, which contradicts (i).

Consider now any list $\mathcal{P}', \mathcal{P}''$, … of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$. Denote $\mathcal{P}' \cup \mathcal{P}'' \cup \cdots$ by $\mathcal{P}$. Trivially, $\mathcal{P} \underset{\text{CED}}{\Rightarrow} (\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots)$ holds. It follows from Lemma 5 that $(\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots) \underset{\text{CSD}}{\Rightarrow} \mathcal{P}$ holds. Together with (40), this gives that

$$(\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots) \underset{\text{CED}}{\Rightarrow} \mathcal{Q} \text{ implies } (\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots) \underset{\text{CSD}}{\Rightarrow} \mathcal{Q}.$$

■

The converse of (40) does not generally hold. Consequently, the attributes of CEDs that can be described in terms of the six properties of mechanisms do not all correspond to attributes

[20] The definition of the conjunction of three or more attributes is the natural extension of the definition for two attributes.

of CSDs. In other words, the former are not simply the restrictions of the latter to correlated equilibrium distributions. Rather, restriction is followed by refinement, which gives rise to additional attributes. Some of the attributes of CEDs, including the majority of those inherited from CSDs, are presented in the following subsection. The subsequent subsection describes additional attributes, by specifically identifying all instances in which the converse of (40) does not hold. The last subsection completes the description of the implication relation $\underset{\text{CED}}{\Rightarrow}$ (henceforth written simply as ⇒) by considering implications involving conjunctions of attributes of CEDs.

5.1 Equivalences

The following propositions identify equivalent formulations for several attributes of CEDs.

Proposition 12. For CEDs, $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$.

Proof. In view of Lemma 4, it suffices to show that if the canonical mechanism $\boldsymbol{m}$ of a CED η has property $\tilde{O}$, then the correlated strategy $\bar{\sigma}$ with the mechanism $\vec{\boldsymbol{m}}$ constructed in the proof of Proposition 5 is a correlated equilibrium. As shown in that proof, if the canonical mechanism satisfies $\tilde{O}$, then $\vec{\boldsymbol{m}}$ satisfies S and $\tilde{O}$.

By Lemma 3, condition (11) is satisfied by the random type profile $\boldsymbol{t}$ and the canonical random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$. By (37), $\bar{\sigma}$ is a correlated equilibrium if and only if

$$E\big(u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i}))\big|\boldsymbol{t}_i, \vec{\boldsymbol{m}}_i(\boldsymbol{t})\big) \geq 0, \qquad i \in N, a_i' \in A_i. \tag{41}$$

Therefore, a sufficient condition for $\bar{\sigma}$ to be a correlated equilibrium is that the conditional expectations in (11) and (41) are equal. The formal difference between them is that player i's action $\boldsymbol{a}_i = \boldsymbol{m}_i(\boldsymbol{t})$ in the former is replaced in the latter by $\vec{\boldsymbol{m}}_i(\boldsymbol{t})$, which by (36) specifies not only the message that player i receives from the canonical mechanism (which is $\boldsymbol{m}_i(\boldsymbol{t})$) but also the messages he would receive if his type were different. Therefore, the meaning of the above equality is that these messages do not provide player i with any information that he could use for choosing a better action.

Since $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$, if the conditional expectations in (11) and (41) were not equal, then by (36) there would be some type of player i, say the first one t_i^1, and some messages $m_i^1, m_i^2, \ldots$ such that

$$E(u_i(\boldsymbol{t}, \boldsymbol{m}(t_i^1, \boldsymbol{t}_{-i})) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{m}_{-i}(t_i^1, \boldsymbol{t}_{-i})))|\boldsymbol{t}_i = t_i^1, \boldsymbol{m}_i(t_i^1, \boldsymbol{t}_{-i}) = m_i^1)$$
$$\neq E(u_i(\boldsymbol{t}, \boldsymbol{m}(t_i^1, \boldsymbol{t}_{-i})) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{m}_{-i}(t_i^1, \boldsymbol{t}_{-i})))|\boldsymbol{t}_i = t_i^1, \boldsymbol{m}_i(t_i^1, \boldsymbol{t}_{-i}) = m_i^1, \boldsymbol{m}_i(t_i^2, \boldsymbol{t}_{-i}) = m_i^2, \ldots).$$

The inequality implies that the pair of random variables $\boldsymbol{t}$ and $\boldsymbol{m}(t_i^1, \boldsymbol{t}_{-i})$ is not independent of $\boldsymbol{m}_i(t_i^2, \boldsymbol{t}_{-i}), \boldsymbol{m}_i(t_i^3, \boldsymbol{t}_{-i}), \ldots$. However, if the canonical mechanism $\boldsymbol{m}$ has property $\tilde{O}$, then is follows from (6) that such independence does in fact hold, so that the above inequality cannot hold, which proves that $\bar{\sigma}$ is a correlated equilibrium. ∎

Proposition 13. For CEDs, $\{\tilde{S}, O\} \Leftrightarrow \{O\}$.

Proof. It suffices to show that for every correlated equilibrium σ with a mechanism $\boldsymbol{m}$ that satisfies O there is another correlated equilibrium $\hat{\sigma}$ with a mechanism $\widehat{\boldsymbol{m}}$ that satisfies $\tilde{S}$

and O such that the two correlated equilibria have identical CEDs. The correlated equilibrium $\hat{\sigma}$ and the mechanism $\widehat{\boldsymbol{m}}$ are constructed below. The construction uses a random variable $\boldsymbol{r}$ that is uniformly distributed on the half-open interval $(0,1]$ and is independent of $\boldsymbol{m}$. (The assumption of uniform distribution, which is inconsistent with the definition of random variable in footnote 2, is only temporary. Below, $\boldsymbol{r}$ is replaced by a random variable that is defined on a finite probability space.)

The idea of the proof is to encode the messages that the mechanism $\boldsymbol{m}$ sends in a particular way. Suppose, without loss of generality, that these messages are integers, more specifically, that the (finite) message space of each player i has the form $M_i = \{1,2,\dots\}$. Since property O of the mechanism implies $\tilde{O}$, the (random) message $\boldsymbol{m}_i(t)$ to player i has a distribution function F_{i,t_i} that only depends on the player's own type t_i. Thus, for any t_{-i},

$$F_{i,t_i}(s) = \Pr(\boldsymbol{m}_i(t_i, t_{-i}) \leq s)\,, \qquad -\infty < s < \infty.$$

The mechanism $\widehat{\boldsymbol{m}}$ combines the message $\boldsymbol{m}_i(t)$ and the random variable $\boldsymbol{r}$ in the following manner:

$$\widehat{\boldsymbol{m}}_i(t) = \boldsymbol{r}\, F_{i,t_i}(\boldsymbol{m}_i(t)) + (1-\boldsymbol{r}) F_{i,t_i}(\boldsymbol{m}_i(t) - 1). \tag{42}$$

It is not difficult to see that the random variable $\widehat{\boldsymbol{m}}_i(t)$ is uniformly distributed on the unit interval for every i and t. Therefore, the mechanism $\widehat{\boldsymbol{m}} = (\widehat{\boldsymbol{m}}(t))_{t\in T}$ satisfies $\tilde{S}$ as well as O. Note that, strictly speaking, the above construction does not conform to the definition of mechanism since the message spaces are infinite. A variant that does not have this problem is described below.

The next step is to define the correlated strategy $\hat{\sigma}$. For each player i, $\hat{\sigma}_i$ is defined by $\hat{\sigma}_i(t_i, \widehat{m}_i) = \sigma_i(t_i, \psi_i(t_i, \widehat{m}_i))$, where ψ_i is a function that "decodes" the message $\widehat{\boldsymbol{m}}_i(t)$ and recovers the original message $\boldsymbol{m}_i(t)$:

$$\psi_i(t_i, x) = \min\{ m_i \in M_i \mid F_{i,t_i}(m_i) \geq x \}.$$

By virtue of this decoding, $\hat{\sigma}_i$ always specifies the same action as σ_i. Since, in addition, the messages that the players receive from the mechanism $\widehat{\boldsymbol{m}}$ convey precisely the same information as those from $\boldsymbol{m}$, this proves that $\hat{\sigma}$, like σ, is a correlated equilibrium.

It remains to replace the uniformly-distribution random variable $\boldsymbol{r}$ with one that has only finitely many possible values, specifically, with the random variable $\mu(\boldsymbol{r})$, where μ is a real-valued function with a finite range. The first step is to consider the (finite) set

$$X = \{F_{i,t_i}(m_i)\}_{i,t_i,m_i}$$

of all values that may appear in the first term in (42). The next step is to modify the definition of the mechanism $\widehat{\boldsymbol{m}}$ by changing the message that it sends to each player i from $\widehat{\boldsymbol{m}}_i(t)$ (which is defined by (42)) to $\lambda(\widehat{\boldsymbol{m}}_i(t))$, where $\lambda\colon [0,1] \to [0,1]$ is the left continuous function defined by $\lambda(x) = \min(\{x' \in X | x' \geq x\})$. This change is inconsequential. Since always $\widehat{\boldsymbol{m}}_i(t) \leq \lambda(\widehat{\boldsymbol{m}}_i(t)) \leq F_{i,t_i}(\boldsymbol{m}_i(t))$, applying the decoder ψ_i to the modified message $\lambda(\widehat{\boldsymbol{m}}_i(t))$ still recovers $\boldsymbol{m}_i(t)$. Let the function $\mu\colon (0,1] \to (0,1]$ be defined by

$$\mu(r) = \max\left\{ 0 < r' \le 1 \;\middle|\; \begin{array}{c} \lambda(r'\, F_{i,t_i}(m_i) + (1-r')F_{i,t_i}(m_i - 1)) \\ = \lambda(r\, F_{i,t_i}(m_i) + (1-r)F_{i,t_i}(m_i - 1)) \\ \text{for all } i, t_i, m_i \end{array} \right\}.$$

It is not difficult to see that the function μ is well defined and, as required, has only finitely many possible values. The final step is to replace $\boldsymbol{r}$ in (42) with $\mu(\boldsymbol{r})$. By definition of μ, this replacement does not change the message $\lambda(\widehat{\boldsymbol{m}}_i(t))$. ■

Proposition 14. For CEDs, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\}$.

Proof. To prove that $\{\tilde{S}, I\} \Rightarrow \{S, I\}$, it has to be shown that every CED implementable by a mechanism $\boldsymbol{m}$ with properties $\tilde{S}$ and I is also implementable by a mechanism with properties S and I.

Property $\tilde{S}$ of $\boldsymbol{m}$ means that for every type profile t and player i the distribution of $\boldsymbol{m}_i(t)$ does not change when only player i's type t_i changes. In other words, the distribution only depends on i and the partial type profile t_{-i}. Therefore, it is possible to construct a family $\{\boldsymbol{r}^{i,t_{-i}}\}$ of independent random variables, indexed by the players and partial type profiles, such that each entry $\boldsymbol{r}^{i,t_{-i}}$ has the distribution described above. For each t, define

$$\widetilde{\boldsymbol{m}}(t) = (\boldsymbol{r}^{1,t_{-1}}, \boldsymbol{r}^{2,t_{-2}}, \dots, \boldsymbol{r}^{n,t_{-n}}).$$

Thus,

$$\widetilde{\boldsymbol{m}}_i(t) \overset{d}{=} \boldsymbol{m}_i(t), \qquad i \in N, t \in T. \tag{43}$$

The mechanism $\widetilde{\boldsymbol{m}} = (\widetilde{\boldsymbol{m}}(t))_{t \in T}$ has properties S and I by construction. Since $\boldsymbol{m}$ also has property I, it follows from (43) that

$$\widetilde{\boldsymbol{m}}(t) \overset{d}{=} \boldsymbol{m}(t), \qquad t \in T, \tag{44}$$

which shows that the two mechanisms are functionally indistinguishable. It is not difficult to see that a correlated strategy with one mechanism is a correlated equilibrium if and only if the same correlated strategy is a correlated equilibrium with the other mechanism. Therefore, the two mechanisms implement precisely the same CEDs.

An almost identical proof shows that $\{\tilde{S}, D\} \Rightarrow \{S, D\}$. The only required change it to assume that the mechanism $\boldsymbol{m}$ has properties $\tilde{S}$ and D. This assumption implies that for every t the distribution of $\boldsymbol{m}(t)$ is degenerate, which by (44) implies the same for $\widetilde{\boldsymbol{m}}(t)$. Thus, $\widetilde{\boldsymbol{m}}$ satisfies D. ■

Proposition 15. For CEDs, $\{S, O, I\} \Leftrightarrow \{\tilde{S}, O, I\} \Leftrightarrow \{S, \tilde{O}, I\} \Leftrightarrow \{\tilde{S}, \tilde{O}, I\} \Leftrightarrow \{O, I\} \Leftrightarrow \{\tilde{O}, I\} \Leftrightarrow (\{\tilde{O}\} \wedge \{I\})$ and $\{S, O, D\} \Leftrightarrow \{\tilde{S}, O, D\} \Leftrightarrow \{S, \tilde{O}, D\} \Leftrightarrow \{\tilde{S}, \tilde{O}, D\} \Leftrightarrow \{O, D\} \Leftrightarrow \{\tilde{O}, D\} \Leftrightarrow (\{\tilde{O}\} \wedge \{D\})$.

Proof. In view of Lemma 4, it suffices to show that if the canonical mechanism $\boldsymbol{m}$ of a CED η has property $\tilde{O}$ as well as I or D, then η is S, O, I- or S, O, D-implementable, respectively. It is shown by Proposition 6 that, as a CSD, η is indeed thus implementable. The proof of that proposition refers to the proof of Proposition 1, where it is shown that the mechanism $\overline{\boldsymbol{m}}$

defined by (23) has the relevant three properties, and with that mechanism, the correlated strategy $\bar{\sigma}$ defined in (26) gives η. Therefore, it only remains to show that the correlated strategy $\bar{\sigma}$ with the mechanism $\overline{\boldsymbol{m}}$ is in fact a correlated equilibrium.

As shown in proof of Proposition 12, $\bar{\sigma}$ is a correlated equilibrium with the mechanism $\overrightarrow{\boldsymbol{m}}$ defined by (36). This mechanism and $\overline{\boldsymbol{m}}$ are both based on the canonical mechanism $\boldsymbol{m}$, and the second can be obtained from the first by selecting a particular type profile t^1 and setting

$$\overline{\boldsymbol{m}}(t) = \overrightarrow{\boldsymbol{m}}(t^1), \qquad t \in T. \tag{45}$$

Therefore, to complete the proof it suffices to show that

$$\overline{\boldsymbol{m}}(t) \overset{d}{=} \overrightarrow{\boldsymbol{m}}(t), \qquad t \in T,$$

so that $\overline{\boldsymbol{m}}$ and $\overrightarrow{\boldsymbol{m}}$ are functionally indistinguishable. To show this, it suffices to establish that replacing t^1 in (45) with any other type profile t' would not chance the distribution of $\overline{\boldsymbol{m}}(t)$. By (25), a sufficient condition for this invariance is that, for all i, t_i and a_i

$$\Pr\left(\boldsymbol{m}_i(t_i, t_{-i}^1) = a_i\right) = \Pr(\boldsymbol{m}_i(t_i, t_{-i}') = a_i).$$

This condition holds since, by assumption, $\boldsymbol{m}$ has property $\tilde{O}$. ■

5.2 Implications

By Proposition 11, an implication relation that does not hold for CSDs also does not hold for CEDs. Therefore, an immediate corollary of Propositions 7, 8, 9 and 10 is the following result.

Proposition 16. For CEDs, $\{S, O, I\} \nRightarrow \{D\}$, $\{S, D\} \nRightarrow \{\tilde{O}\}$, $\{\tilde{O}\} \nRightarrow \{O\}$ and $\{S, O\} \nRightarrow \{I\}$.

The next three propositions identify implication relations that do not hold for CEDs even though they hold for CSDs.

Proposition 17. For CEDs, $\{D\} \nRightarrow \{\tilde{S}\}$.

Proof. This is proved by Example 2. ■

Proposition 18. For CEDs, $\{\tilde{S}\} \nRightarrow \{S\}$.

Proof. This is proved by Example 3. ■

Proposition 19. For CEDs, $\{\tilde{S}, O\} \nRightarrow \{S, O\}$.

Proof. This is proved by Example 6. ■

Propositions 12, 13, 14 and 15 identify six attributes of correlated equilibrium distributions that are defined by subsets of the properties of mechanisms defined in Section 2.2. These attributes plus S-implementability are shown in Figure 2 as attributes I_a, II, III, IV_a, V, VI_a and VII. The implication relations that are specified by the Hesse diagram among these attributes all hold trivially (since they follow immediately from relations between properties of mechanisms). By Proposition 16, additional implications among the seven attributes do not hold, and in particular, none of them is equivalent to any of the others. (The more detailed argument given in Section 4.3 also applies here, mutatis mutandis.) Three more attributes are defined by $\{\,\}$, $\{\tilde{S}\}$ and $\{V\}$ (I, $\mathrm{I_b}$ and VI in Figure 2). The implication relations shown in Figure 2 among these attributes and between them and the other seven all hold trivially. It follows from Propositions 17 and 18, and from $\{S, O, I\} \nRightarrow \{D\}$ in Proposition 16, that additional such implications do not hold. Two more attributes are defined by $\{S, O\}$ and $\{I\}$. It follows from Proposition 19, and from $\{S, O\} \nRightarrow \{I\}$ in Proposition 16, that the implication relations shown in Figure 2 between each of these attributes and each of the other ten are the only ones holding. This proves that there are precisely twelve distinct attributes of CEDs that can be defined by single subsets of the properties of mechanisms in Section 2.2.

As indicated in Section 3.2.1, attributes I, II, III, IV, V, VI and VII of CEDs are obtained from the similarly numbered attributes of CSDs by restriction. That is, a CED has any of these attributes if and only if it has the corresponding attribute as a CSD. This result is an immediate corollary of Theorem 1, which is stated in Section 3.2.1 and is proved below.

Proof of Theorem 1. Suppose first that $\mathcal{P} \subseteq \{\tilde{O}, D, I\}$, and let η be a CED that is $\mathcal{P}$-implementable as a CSD. By Lemma 4, the canonical mechanism of η has all the properties in $\mathcal{P}$. Therefore, by Corollary 1, η is $\mathcal{P}$-implementable also as a CED.

Next, consider the case $\mathcal{P} = \{O\}$. As remarked in Section 4, for every O-implementable CSD η there is a mechanism $\widetilde{\boldsymbol{m}}$ with property O that is functionally indistinguishable from the canonical mechanism. If η is moreover a CED, then by Corollary 1 the canonical strategy with the canonical mechanism is a correlated equilibrium, and the same is therefore true with the canonical mechanism replaced by $\widetilde{\boldsymbol{m}}$. Hence, η is O-implementable also as a CED.

To complete the proof of the theorem it remains to note that, by Proposition 6 and 15, for both CSDs and CEDs, O, I-implementability and $\tilde{O}, I$-implementability are equivalent, and the same is true for O, D- and $\tilde{O}, D$-implementability. ■

5.3 Conjunction of attributes

The next step is to consider attributes of CEDs that are defined by pairs (or possibly triplets, etc.) of subsets of properties of mechanisms, that is, by conjunction of two (or more) of the twelve attributes identified above. Unlike for CSDs (see Lemma 5), genuinely new attributes can be defined this way. For example, it follows from the second part of the next proposition that the conjunction of S-implementability and D-implementability is a new attribute.

Proposition 20. For CEDs, $(\{S\} \wedge \{D\}) \Leftrightarrow \left(\{\tilde{S}\} \wedge \{D\}\right) \Rightarrow \{S, I\}$ but $(\{S\} \wedge \{D\}) \nRightarrow \{S, D\}$.

Proof. To prove the first part of the proposition, it suffices to show that $\left(\{\tilde{S}\} \wedge \{D\}\right) \Rightarrow \{S, I\}$; the equivalence then follows immediately from the trivial implications $\{S, I\} \Rightarrow \{S\} \Rightarrow \{\tilde{S}\}$.

Consider a CED η that is both $\tilde{S}$- and D-implementable. It has to be shown that η is also S, I-implementable. By the assumption of D-implementability and (condition (i) in) Proposition 4, there is a mapping $\phi = (\phi_1, \phi_2, \dots, \phi_n): T \to A$ such that $\eta(\{(t, \phi(t))\}) = \eta_T(\{t\})$ for all type profiles t. By the assumption of $\tilde{S}$-implementability, there is a correlated strategy σ with a mechanism $\boldsymbol{m}$ satisfying $\tilde{S}$ such that η is equal to the joint distribution of the random type profile $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ defined by (7). In particular, for every $t = (t_1, t_2, \dots, t_n) \in \mathrm{supp}(\eta_T)$,

$$\Pr(\sigma_i(t_i, \boldsymbol{m}_i(t)) = \phi_i(t) \text{ for all } i) = \Pr(\boldsymbol{a} = \phi(t) | \boldsymbol{t} = t) = \frac{\eta\big(\{(t, \phi(t))\}\big)}{\eta_T(\{t\})} = 1.$$

Therefore,

$$\boldsymbol{a}_i = \sigma_i(\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})) = \phi_i(\boldsymbol{t}), \qquad i \in N. \tag{46}$$

Let the mechanism $\widetilde{\boldsymbol{m}}$ be as in the proof of Proposition 14. By (43) and (46),

$$\sigma_i(\boldsymbol{t}_i, \widetilde{\boldsymbol{m}}_i(\boldsymbol{t})) = \phi_i(\boldsymbol{t}), \qquad i \in N.$$

Therefore, using the correlated strategy σ with the mechanism $\widetilde{\boldsymbol{m}}$ instead of $\boldsymbol{m}$ also gives the distribution η. Moreover, if $\widetilde{\boldsymbol{m}}$ is used and a single player i changes his strategy from σ_i to some other strategy σ_i', the player's expected payoff changes to $E(u_i(\boldsymbol{t}, (\sigma_i'(\boldsymbol{t}_i, \widetilde{\boldsymbol{m}}_i(\boldsymbol{t})), \phi_{-i}(\boldsymbol{t}))))$. By (43), this new payoff is equal to

$$E\big(u_i(\boldsymbol{t}, (\sigma_i'(\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})), \phi_{-i}(\boldsymbol{t})))\big),$$

which by (46) is i's expected payoff if he unilaterally changes his strategy from σ_i to σ_i' when the correlated strategy σ is used with the mechanism $\boldsymbol{m}$ (rather than $\widetilde{\boldsymbol{m}}$). Since with this mechanism the correlated strategy is a correlated equilibrium, i's change of strategy cannot increase his expected payoff. This proves that σ is an equilibrium also with the mechanism $\widetilde{\boldsymbol{m}}$, which by construction satisfies S and I.

The second part of the proposition is proved by Example 4.[21] ∎

It follows from the next proposition that the conjunction of S-implementability and I-implementability is a new attribute.

Proposition 21. For CEDs, $(\{S\} \wedge \{I\}) \not\Rightarrow \{S, I\}$.

Proof. This is shown by Example 5. ∎

Whether the conjunction of $\tilde{S}$-implementability and I-implementability is also a new attribute of CEDs is not known. It depends on the answer to the following question.

Open Question. For CEDs, does $\big(\{\tilde{S}\} \wedge \{I\}\big) \Rightarrow \{S\}$?

This question corresponds to the question mark in Figure 2. The marked line in the Hesse diagram exists if and only if the answer is negative, which means that there is a CED in some Bayesian game that is both $\tilde{S}$- and I-implementable but not S-implementable. If the answer is affirmative, the two attributes of CEDs connected by the line ($\mathrm{IV_b}$ and the unnumbered attribute) are actually one and the same, that is, they are equivalent attributes.

[21] Parenthetically, an argument broadly similar to that used above shows that $(\{S\} \wedge \{D\}) \Rightarrow \{S, D\}$ *would* hold if it were assumed that the type distribution η_T has full support, i.e., $\mathrm{supp}(\eta_T) = T$ (so that every type profile has positive probability).

Depending on the answer to the Open Question, there are two or three attributes of CEDs that can be defined as the conjunction of a pair of incomparable attributes of the twelve ones presented above. Thus, there are in total 14 or 15 attributes of CEDs, which are related to one another as in Figure 2. The following lemma shows that this list is complete in that there are no additional, nonequivalent attributes that can be defined as the conjunction of two or more of those in Figure 2. This result holds regardless of the answer to the Open Question.

Lemma 6. The conjunction of any number of the attributes of CEDs in Figure 2 is equivalent to one of the attributes in the same figure.

Proof of Lemma 6 (an outline). Proposition 15 proves the two special cases of the conjunctions of $\tilde{O}$-implementability and either D- or I-implementability. For the general case, it has to be shown that for every list $\mathcal{P}', \mathcal{P}'', \ldots \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$, the conjunction $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$ is equivalent to one of the attributes in Figure 2. (It suffices to consider lists with three or fewer entries, since in any longer list at least two elements represent comparable attributes.) This can be shown quite easily in a straightforward, case-by-case manner. ∎

6 Communication Equilibrium Distributions

As for correlated strategy distributions and correlated equilibrium distributions, different kinds of mechanisms implement different kinds of communication equilibrium distributions. Specifically, for each subset $\mathcal{P}$ of the six properties of mechanisms described in Section 2.2, a MED is $\mathcal{P}$-implementable if it is given by some communication equilibrium with a mechanism that has (all) the properties in $\mathcal{P}$. This section, like the previous two, is mainly concerned with the implication relation between these attributes, and conjunctions of several attributes. Implication is denoted by the generic symbol ⇒ when it is clear from the context that it refers to attributes of MEDs. Otherwise, the more explicit symbol $\underset{\text{MED}}{\Rightarrow}$ is used. As it turns out, a *necessary* condition for the implication to hold is that a similar relation holds for CEDs. The proof of the following proposition is given at the end of this section.

Proposition 22. For every two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$\mathcal{P} \underset{\text{MED}}{\Rightarrow} \mathcal{Q} \;\text{ implies }\; \mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}. \tag{47}$$

Moreover, the same is true with $\mathcal{P}$ replaced by by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$, for any list $\mathcal{P}', \mathcal{P}'', \ldots$ of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$.

Since by definition every MED is also a CED, it may seem that the reverse of the implication (47) also ought to hold. However, as remarked in Section 3.3, Example 7 shows that this is in fact not so, which is why the Hesse diagrams of the implications relations between attributes of CEDs (Figure 2) and between attributes of MEDs (Figure 3) are different. However, the differences only concern conjunction of attributes. As shown below, for attributes that a defined by single sets for properties of mechanisms, the implication relations for CEDs and MEDs are identical.

6.1 Equivalences

The following propositions parallel those in Section 5.1. Thus, they identify instances in which the reverse of (47) also holds.

Proposition 23. For MEDs, $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$.

Proof. It has to be shown that for every communication equilibrium σ with a mechanism $\boldsymbol{m}$ that has property $\tilde{O}$, the resulting MED η is $S, \tilde{O}$-implementable.

Unlike in the proof of Proposition 12, the mechanism $\boldsymbol{m}$ is not necessarily canonical. Nevertheless, without loss of generality, it may be assumed that it satisfies (6). Otherwise, $\boldsymbol{m}$ could be replaced by any mechanism $\widetilde{\boldsymbol{m}}$ satisfying (6) such that

$$\widetilde{\boldsymbol{m}}(t) \stackrel{d}{=} \boldsymbol{m}(t), \qquad t \in T.$$

These equalities imply that $\widetilde{\boldsymbol{m}}$ also has property $\tilde{O}$, and it is not difficult to see that σ with $\widetilde{\boldsymbol{m}}$ is also a communication equilibrium, since for any profile of reported types, the messages that $\widetilde{\boldsymbol{m}}$ sends are indistinguishable from those of $\boldsymbol{m}$.

Consider the correlated strategy $\bar{\sigma}$ and the mechanism $\vec{\boldsymbol{m}} = (\vec{\boldsymbol{m}}(t))_{t \in T}$, which are defined, respectively, by (26) and the following generalization of (36):

$$\vec{\boldsymbol{m}}_i(t) = \left(\sigma_i(t_i^1, \boldsymbol{m}_i(t_i^1, t_{-i})), \sigma_i(t_i^2, \boldsymbol{m}_i(t_i^2, t_{-i})), \ldots\right), \qquad i \in N, t \in T.$$

Arguments similar to those used in the proof of Proposition 5 show that $\vec{\boldsymbol{m}}$ has properties S and $\tilde{O}$. It has to be shown that, with this mechanism, $\bar{\sigma}$ is a communication equilibrium. That is, for every player i, type t_i' for that player and function $\bar{\sigma}_i' : T_i \times A_i^{T_i} \to A_i$,

$$E(u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, \boldsymbol{a}') | \boldsymbol{t}_i) \geq 0, \tag{48}$$

where $\boldsymbol{a}$ is the random action profile corresponding to $\bar{\sigma}$ and $\boldsymbol{a}'$ is obtained from (13) by replacing σ_i', σ_j and $\boldsymbol{m}$ by $\bar{\sigma}_i'$, $\bar{\sigma}_j$ and $\vec{\boldsymbol{m}}$, respectively. Obviously, it suffices to consider the (effectively, completely general) case $i = 1$ and $t_1' = t_1^1$, for which

$$\boldsymbol{a}_j = \bar{\sigma}_j(\boldsymbol{t}_j, \vec{\boldsymbol{m}}_j(\boldsymbol{t})) = \sigma_j(\boldsymbol{t}_j, \boldsymbol{m}_j(\boldsymbol{t})), \qquad j \in N \tag{49}$$

and

$$\begin{aligned} \boldsymbol{a}_1' &= \bar{\sigma}_1'(\boldsymbol{t}_1, \vec{\boldsymbol{m}}_1(t_1^1, \boldsymbol{t}_{-1})) = \bar{\sigma}_1'(\boldsymbol{t}_1, (\sigma_1(t_1^1, \boldsymbol{m}_1(t_1^1, \boldsymbol{t}_{-1})), \sigma_1(t_1^2, \boldsymbol{m}_1(t_1^2, \boldsymbol{t}_{-1})), \ldots)), \\ \boldsymbol{a}_j' &= \bar{\sigma}_j(\boldsymbol{t}_j, \vec{\boldsymbol{m}}_j(t_1^1, \boldsymbol{t}_{-1})) = \sigma_j(\boldsymbol{t}_j, \boldsymbol{m}_j(t_1^1, \boldsymbol{t}_{-1})), \qquad j \neq 1. \end{aligned}$$

If (48) does not hold (for $i = 1$), then there are some t_1'' and $m_1^2, m_1^3, \ldots$ such that

$$E(u_1(\boldsymbol{t}, \boldsymbol{a}') \big| \boldsymbol{t}_1 = t_1'', \boldsymbol{m}_1(t_1^2, \boldsymbol{t}_{-1}) = m_1^2, \boldsymbol{m}_1(t_1^3, \boldsymbol{t}_{-1}) = m_1^3, \ldots) > E(u_1(\boldsymbol{t}, \boldsymbol{a}) | \boldsymbol{t}_1 = t_1'').$$

It follows from properties $\tilde{O}$ and (6) of $\boldsymbol{m}$ that the pair of random variables $\boldsymbol{t}$ and $\boldsymbol{m}(t_1^1, \boldsymbol{t}_{-1})$ is independent of $\boldsymbol{m}_1(t_1^2, \boldsymbol{t}_{-1}), \boldsymbol{m}_1(t_1^3, \boldsymbol{t}_{-1}), \ldots$. Therefore, the above inequality is equivalent to

$$E(u_1(\boldsymbol{t}, (\sigma_1'(\boldsymbol{t}_1, \boldsymbol{m}_1(t_1^1, \boldsymbol{t}_{-1})), \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(t_1^1, \boldsymbol{t}_{-1})), \dots \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(t_1^1, \boldsymbol{t}_{-1}))))\big|\boldsymbol{t}_1 = t_1'') \qquad (50)$$
$$> E(u_1(\boldsymbol{t}, \boldsymbol{a})|\boldsymbol{t}_1 = t_1''),$$

where $\sigma_1': T_1 \times M_1 \to A_1$ is the function defined by

$$\sigma_1'(t_1, m_1) = \bar{\sigma}_1'(t_1, (\sigma_1(t_1^1, m_1), \sigma_1(t_1^2, m_1^2), \sigma_1(t_1^3, m_1^3), \dots)).$$

However, in conjunction with (49), inequality (50) contradicts the assumption that σ with the mechanism $\boldsymbol{m}$ is a communication equilibrium, since it shows that when player 1's type is t_1'', he can gain from misreporting it as t_1^1 and switching from σ_1 to σ_1'. The contradiction proves that the correlated strategy $\bar{\sigma}$ with the mechanism $\vec{\boldsymbol{m}}$ is a communication equilibrium. ■

Proposition 24. For MEDs, $\{\tilde{S}, O\} \Leftrightarrow \{O\}$.

Proof. Identical to the proof of Proposition 13. ■

Proposition 25. For MEDs, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\}$.

Proof. Identical to the proof of Proposition 14. ■

Proposition 26. For MEDs, $\{S, O, I\} \Leftrightarrow \{\tilde{S}, O, I\} \Leftrightarrow \{S, \tilde{O}, I\} \Leftrightarrow \{\tilde{S}, \tilde{O}, I\} \Leftrightarrow \{O, I\} \Leftrightarrow \{\tilde{O}, I\} \Leftrightarrow (\{\tilde{O}\} \wedge \{I\})$ and $\{S, O, D\} \Leftrightarrow \{\tilde{S}, O, D\} \Leftrightarrow \{S, \tilde{O}, D\} \Leftrightarrow \{\tilde{S}, \tilde{O}, D\} \Leftrightarrow \{O, D\} \Leftrightarrow \{\tilde{O}, D\} \Leftrightarrow (\{\tilde{O}\} \wedge \{D\})$.

Proof. It suffices to show that $(\{\tilde{O}\} \wedge \{I\}) \Rightarrow \{S, O, I\}$, and similarly with I replaced by D. By Proposition 15, both implications hold of CEDs, and the result that they also hold for MEDs follows immediately from the fact that a S, O, I-implementable CED is automatically a MED. ■

6.2 Implications

Implication (47) in Proposition 22 can equivalently be expressed by its counterpositive: If a counterexample of a $\mathcal{P}$-implementable CED that is not $\mathcal{Q}$-implementable exists, then a similar counterexample can be found for MEDs. Finding the latter can be easy or quite complicated. The former holds if the CED example employs a correlated equilibrium (with a mechanism with the properties in $\mathcal{P}$) that is also a communication equilibrium: players have no incentive to lie about their types. In this case, the same example can be used for MEDs, since a CED that is not $\mathcal{Q}$-implementable a fortiori does not have that attribute as a MED. The proofs of the following two propositions use this simple observation.

Proposition 27. For MEDs, $\{\tilde{O}\} \nRightarrow \{O\}$, $\{S, O\} \nRightarrow \{I\}$ $\{S, D\} \nRightarrow \{\tilde{O}\}$ and $\{S, O, I\} \nRightarrow \{D\}$.

Proof. Proposition 16, which establishes the same for CEDs, relies on Proposition 11. Therefore, it suffices to show that a result similar to the latter holds with correlated equilibrium (distribution) replaced with communication equilibrium (distribution). This can be shown by simply making this replacement throughout the proof of Proposition 11. ■

Proposition 28. For MEDs, $\{D\} \nRightarrow \{\tilde{S}\}$.

Proof. The correlated equilibrium with the mechanism with property D that is described in Example 2 is in fact a communication equilibrium. If player 2 lies about his type, player 1 will get as a message an incorrect type profile and will consequentially choose an action for which a positive payoff for 2 is impossible. For a similar reason, player 1 cannot gain from lying; in this case, the lie will only affect type $+1$ of player 2. The MED of this communication equilibrium in not $\tilde{S}$-implementable since, as shown, it does not have this attribute even as a CED. ■

Even if the correlated equilibrium that proves that a certain implication does not hold for CEDs is not a communication equilibrium, it may be possible to make truthful type reports incentive compatible by augmenting the original game with a suitable auxiliary game and modifying the mechanism and correlated strategy accordingly.

Suppose, for example, that each of the two players in a Bayesian game can have type $+1$ or -1, and all four type profiles are equally probable. The game can then be modified by adding to it an auxiliary game that requires each player to push one of three buttons, B_1, B_2 or B_3. Depending on both players' choices of button and on their types, a very large number $K > 0$ is then either added to or subtracted from their payoffs in the original game. Specifically, the change in payoffs ($+K$ or $-K$) is determined according to the following table, where the rows and columns correspond to the choices of player 1 and 2, respectively, and $\tau = t_1 t_2$ is the product of their types:

	B_1	B_2	B_3
B_1	τK	$-\tau K$	$-K$
B_2	$-K$	τK	$-\tau K$
B_3	$-\tau K$	$-K$	τK

Thus, both for $\tau = +1$ and for -1, three cells in the table represent reward and six represent punishment. Any mechanism in the original game can be turned into one in the augmented game by appending to the message it sends to each player, which pertains to the original game, a recommendation of button in the auxiliary game. The latter is determined in the following way. The mechanism attempts to identify the "rewarding" cells by calculating the product of the players' *reported* types, it randomly selects one of these cells (each with probability $1/3$), and it recommends its row and column to player 1 and 2, respectively. As detailed below, the feature of the mechanism that encourages truth telling is that misreporting will result in misidentification of the rewarding cells. Note that, for any pair of (reported) types, the recommendation to each player is equally likely to be B_1, B_2 or B_3. Therefore, the modified mechanism has property $\tilde{S}$ or $\tilde{O}$ if the original mechanism has the same property. It cannot, however, have any of the other four properties of mechanisms. (With a somewhat more sophisticated auxiliary game, it is possible to also retain property S.)

To any correlated equilibrium in the original game there corresponds a communication equilibrium in the augmented game. In that equilibrium, the mechanism appends recommendations as described above, each player pushes the recommended button, and plays in the original game according to the original correlated equilibrium. To see that truthful type reports are incentive compatible, suppose that, for example, button B_1 is recommended to type $+1$ of player 1. If both players reported their types truthfully and they follow the recommendations of the mechanism, the player can infer that player 2 will choose B_1 or B_2 if his type is $+1$ or -1, respectively, and in both cases, the players will get the reward K. However, if (only) player 1 misreported his type, then player 2's type has the opposite relation with his action. Player 1 must then choose B_2 or B_3 if 2's type is $+1$ or -1, respectively; otherwise, K will be subtracted from rather than added to his payoff. Since the players' types are independent, this means that he cannot get more than zero in expectation. Thus, dishonesty does not pay.

Proposition 29. For MEDs, $\{\tilde{S}\} \not\Rightarrow \{S\}$.

Proof. Consider the Bayesian game and the $\tilde{S}$- but not S-implementable CED presented in Example 3. This CED is not a MED. However, a communication equilibrium with a mechanism that has property $\tilde{S}$ can be obtained by modifying the game and the correlated equilibrium described in the example by adding an auxiliary game as above. The corresponding MED is not S-implementable even as CED. It is not difficult to see that, if it were S-implementable, the same would be true for the original CED. ■

The proofs of the next two propositions involve more particular modifications of the original counterexamples (that is, those refereeing to CEDs).

Proposition 30. For MEDs, $\{\tilde{S}, O\} \not\Rightarrow \{S, O\}$.

Proof. Consider the following changes to the game and CED in Example 6. Both players can have type $+1$ or -1, and all four type profiles are equally probable. If the players' types differ or are identical, they both receive the payoff specified by the matrix

$$\begin{matrix} & L & R \\ L & \multirow{2}{*}{$\begin{pmatrix} 0 & 0 \\ 0 & 1.5 \end{pmatrix}$} \\ R & \end{matrix} \quad \text{or} \quad \begin{matrix} & L & R \\ L & \multirow{2}{*}{$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$} \\ R & \end{matrix},$$

respectively. With the mechanism described in Example 6, the correlated strategy of following the mechanism's recommendations is still a correlated equilibrium. For a player of any type who receives the message L and takes that action, the expected payoff is $1/2 \cdot 0 + 1/2 \cdot 1 = 0.5$, whereas playing R instead would only yield $1/2 \cdot 1/2 \cdot 1.5 + 1/2 \cdot 0 = 0.375$. If the message is R, taking that action gives 0.875 while playing L would give 0. This correlated equilibrium is moreover a communication equilibrium. If a player misreports his type, he will maximize his payoff by taking the recommended action, since this is also the action that the other player will take if the (real) types differ (and if the types are identical, then the expected payoff from any action is 0.5). Thus, a dishonest player cannot get more than $1/2 \cdot 1/2 \cdot 1.5 + 1/2 \cdot 0.5 = 0.625$, which is less than the $1/2 \cdot 0.5 + 1/2 \cdot 0.875 = 0.6875$ a truthful report would yield.

It remains to show that the corresponding MED is different from that given by any communication (or even correlated) equilibrium σ with a mechanism that has properties S and O. The messages that such a mechanism sends to the players can be written as $\boldsymbol{m}(t')$, for arbitrary type profile t'. Since the players' actions are identical if their types are identical, necessarily

$$(\sigma_1(+1, \boldsymbol{m}_1(t')), \sigma_1(-1, \boldsymbol{m}_1(t')) = (\sigma_2(+1, \boldsymbol{m}_2(t')), \sigma_2(-1, \boldsymbol{m}_2(t')). \quad (51)$$

If $\boldsymbol{m}_1(t')$ is such that the left- (and, hence, also the right-) hand side equals (L, R) or (R, L), respectively, then type $+1$ or -1 of player 1 will get $1/2 \cdot 1$ from taking the action L he is supposed to take but $1/2 \cdot 1.5$ from playing R. Therefore, with probability 1 all four actions in (51) must be identical, which shows that the above MED, in which the players' actions may differ, cannot be obtained. ■

Proposition 31. For MEDs, $(\{S\} \wedge \{D\}) \nRightarrow \{S, I\}$ and $(\{S, I\} \wedge \{D\}) \nRightarrow \{S, D\}$.

Proof. The first part is proved by Example 7. To prove the second part, consider the following changes to the game and CED in Example 4. Player 2 has the constant payoff 0, and he is allowed to choose action R as well as L. Choosing R rather than L reduces by 3 the payoff of type t_1'' of player 1, and has no effect on the other payoffs. The two mechanisms considered in Example 4 and the corresponding correlated equilibria are modified as follows. Both mechanisms instruct player 2 to play R if player 1 reports the type t_1' and to play L otherwise, and player 2 obeys. Clearly, this means that type t_1'' of player 1 has an incentive to report his type truthfully. The same is true for type t_1', for whom the CED gives the highest possible payoff. ■

An alternative proof for the last proposition could be obtained by using the following simple and generally applicable modification of the game and correlated equilibria in the original example. Instead of changing the players' action spaces, a new player is added to the game. This "player 0" has a single type, and his action space is the collection T of type profiles of the original players. If the action he chooses coincides with the original players' actual type profile, everyone gets a huge bonus. Any correlated equilibrium in the original game can be modified as follows. The mechanism sends to player 0 the type reports of the other players, and he chooses the corresponding action. This obviously creates an incentive for the players to report their types truthfully, and thus turns the correlated equilibrium into a communication equilibrium (in the modified game). If property S, $\tilde{S}$, I or D holds for the original mechanism, the modified one also has the same property.

The following proposition uses this construction to show that if the answer to the Open Question presented in Section 5.3 is negative, then the same is true for MEDs. Note that if the answer will turn out to be affirmative, the proposition is uninformative, since its assertion holds vacuously.

Proposition 32. If, for MEDs, $(\{\tilde{S}\} \wedge \{I\}) \Rightarrow \{S\}$, then the same is true for CEDs

Proof. Suppose there is a CED η in some Bayesian game which is not S-implementable but it is given by some correlated equilibrium σ with a mechanism $\boldsymbol{m}$ that has property $\tilde{S}$ as well as by a correlated equilibrium σ' with a mechanism $\boldsymbol{m}'$ that has property I. It has to be shown that a MED with similar properties also exists.

Modify the game and the two correlated equilibria that give η by adding a new player, as detailed above, thus turning σ and σ' into communication equilibria, whose MEDs are identical. The MED assigns nonzero probability only to pairs of type and action profiles in which the former coincides with the action of player 0, and the probability in this case is equal to that assigned by η to the pair obtained by omitting player 0's action. Any communication, or even just correlated, equilibrium that gives this MED can be turned into a correlated equilibrium in the original game (in which η is a CED) simply by omitting the message to player 0 and the corresponding coordinate of the correlated strategy. If the mechanism had property S, the omission would not affect it. It therefore follows from the assumption concerning the CED η that a mechanism implementing the MED cannot in fact have property S. ■

Propositions 23, 24, 25 and 26 identify six attributes of communication equilibrium distributions that are defined by subsets of the properties of mechanisms described in Section 2.2. Figure 3 shows these attributes, marked II, III, IV_a, V, VI_a and VII, and eleven additional ones. The implication relations that are specified by the Hesse diagram among these 17 attributes all hold trivially, since they follow immediately from relations between properties of mechanisms. For two of the implications, it is not known whether the inverse implication also holds. The uncertainty is indicated in Figure 3 by a question mark. If the inverse implication does hold, then the marked line should be removed and the two connected boxes should be coalesced, as they represent equivalent attributes. The following arguments show that none of the other attributes in Figure 3 are equivalent, and more generally, that the Hesse diagram shows *all* the implication relations between attributes of MEDs.

If attributes that involve conjunctions were removed from Figure 2 and Figure 3, the two Hesse diagrams would become identical. In Section 5 it is shown that, among the remaining twelve attributes of CEDs, the implications shown in the diagram are the only ones holding. Essentially the same arguments prove the same for MEDs, with Propositions 27, 28, 29 and 30 replacing 16, 17, 18 and 19, respectively. For each of the attributes in Figure 3 that does involve conjunction, it follows from Proposition 31 that the only other attributes that imply or are implied by it are those indicated as such by the Hesse diagram. This proves that the diagram is complete in terms of implication relations. The following lemma shows that it is also complete in terms of closedness under conjunctions.

Lemma 7. The conjunction of any number of the attributes of MEDs in Figure 3 is equivalent to one of the attributes in the same figure.

Proof of Lemma 7. The proof is similar to that of Lemma 6, except that is uses Proposition 26 instead of 15. ■

It is now possible to show that the implication relation $\underset{\text{MED}}{\Rightarrow}$ between attributes of MEDs is indeed stronger, in a sense, than the corresponding relation for CEDs.

Proof of Proposition 22. It suffices to consider $\mathcal{P}$ and $\mathcal{Q}$ that belong to the group of 27 subsets shown in Figure 2 and Figure 3 (see Section 1). As indicated, if attributes that involve conjunctions are removed from these diagrams, they become identical. This proves the implication (47) as well as the reverse one.

For the more general version in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}''$, it suffices to consider the case in which $\mathcal{P}'$- and $\mathcal{P}''$-implementability (of CEDs, or equivalently MEDs) are incomparable; the case in which they are comparable reduces to the version just analyzed. Consider the meet (greatest lower bound) of $\mathcal{P}'$- and $\mathcal{P}''$-implementability in Figure 3, and their meet in Figure 2. The two are not necessarily similar in any sense. The former is an attribute of MEDs and the latter an attribute of CEDs. However, an exhaustive straightforward examination shows that, with a single exception, for all $\mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$, if the first attribute implies $\mathcal{Q}$-implementability in Figure 3, then the second one does so in Figure 2 (but not necessarily conversely). The single exception is the case covered by Proposition 32: $\mathcal{P}'$ and $\mathcal{P}''$ are $\{\tilde{S}\}$ and $\{I\}$, and $\mathcal{Q} = \{S\}$. This proves the version of (47) in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}''$.

For the version in which the list $\mathcal{P}', \mathcal{P}''$, ... has three or more elements, it again suffices to consider the case in which no two elements describe comparable attributes. It is, however, not difficult to check that this means that, in both diagrams, $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$ is (equivalent to) attribute VII. Therefore, the version of (47) in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$ holds trivially. ∎

7 Correlated Strategy, Correlated Equilibrium and Communication Equilibrium Payoffs

The expected payoffs of the players in a Bayesian game are completely determined by the joint distribution of the players' types and actions. However, the relation between distributions and payoff vectors is normally many-to-one. Therefore, if a particular correlated equilibrium distribution cannot be implemented by a mechanism of a particular kind, it does not necessarily follow that the corresponding payoff vector is not implementable; it may be that a mechanism of that kind implements another CED with the same payoffs. Games with constant payoff functions provide a trivial example of this. In such games, a CED is implementable if and only if it is implementable as a CSD, so that the connection between implementability of distributions and the properties of the mechanism is as detailed in Section 4. By contrast, the single possible payoff vector is of course implementable by any mechanism.

Correlated strategy payoffs (CSPs), correlated equilibrium payoffs (CEPs) and communication equilibrium payoffs (MEPs) in Bayesian games can be classified in a manner similar to the classification of CSDs (Figure 1), CEDs (Figure 2) and MEDs (Figure 3). Each subset $\mathcal{P}$ of the properties of mechanisms described in Section 2.2 defines an attribute of CSPs, CEPs and MEPs, namely, $\mathcal{P}$-implementability. A payoff vector $v = (v_1, v_2, \dots, v_n) \in \mathfrak{R}^n$ in a specified n-player Bayesian game is $\mathcal{P}$-implementable if it is obtained in some correlated strategy, correlated equilibrium or communication equilibrium in the game with a mechanism that has (all) the properties in $\mathcal{P}$ (equivalently, if it is obtained in some CSD, CED or MED, respectively, that is implementable by such a mechanism). For two subset $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$, $\mathcal{P}$-implementability of CSPs implies $\mathcal{Q}$-implementability if in every Bayesian game every CSP that is implementable by some mechanism with the properties in $\mathcal{P}$ is also implementable by a mechanism with the properties in $\mathcal{Q}$. This relation is written as $\mathcal{P} \underset{\text{CSP}}{\Rightarrow} \mathcal{Q}$. For CEPs and MEPs, the relations $\mathcal{P} \underset{\text{CEP}}{\Rightarrow} \mathcal{Q}$ and $\mathcal{P} \underset{\text{MEP}}{\Rightarrow} \mathcal{Q}$ are defined in a similar way.

The main result concerning implementability of payoff vectors is Theorem 4 (Section 3.5). According to this theorem, which is proved below, both aspects of the equilibrium outcomes, the joint distributions of types and actions and the resulting payoffs, are affected by the properties of the implementing mechanisms in a similar way. More precisely, there is a one-to-one correspondence between attributes of CSDs and attributes of CSPs, which are both described by the Hesse diagram in Figure 1, and similar correspondences exists between attributes of CEDs and CEPs (Figure 2), and between attributes of MEDs and MEPs (Figure 3). Note that the example of a game with constant payoffs does not contradict these findings. The properties of the implementing mechanism need only limit the equilibrium payoffs in *some* Bayesian game. They do not have to (and they cannot) come into play in every game.

Proof of Theorem 4. The proofs for correlated equilibria and for communication equilibria are nearly identical. Only the former is presented below; the latter can be obtained from it essentially by replacing 'correlated' with 'communication' throughout. The proof for correlated strategies can also be easily obtained from the proof below by simplifying it in the obvious manner.

It has to be shown that, for every $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$\mathcal{P} \underset{\text{CEP}}{\Rightarrow} \mathcal{Q} \text{ if and only if } \mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}, \tag{52}$$

and that the same is true with $\mathcal{P}$ replaced by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$, for any list $\mathcal{P}', \mathcal{P}''$, ... of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$. One direction of (52) ("if") is easy. $\mathcal{P} \underset{\text{CEP}}{\not\Rightarrow} \mathcal{Q}$ and $\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}$ cannot both hold, since the former means that, in some Bayesian game, there is a $\mathcal{P}$-implementable CED η with a payoff vector that is different from that of every $\mathcal{Q}$-implementable CED in the same game, whereas the latter implies that η itself is $\mathcal{Q}$-implementable.

To prove the nontrivial direction of (52) ("only if"), define the *extension* of a Bayesian game as the game obtained by the addition of *dummy players* — one for each element of $T \times A$. A dummy player has only one possible type and one action, which are therefore insignificant in that they cannot affect the payoff of any player. In the following, the types and actions of the dummy players are ignored, and the collections of type profiles and action profiles in the extended game are thus identified with those in the original game (namely, T and A, respectively). The significance of the dummy players lies in their payoff functions. The payoff function $u_{t,a}: T \times A \to \mathfrak{R}$ of the dummy player representing the types-actions pair $(t, a) \in T \times A$ is defined as the indicator function $1_{\{(t,a)\}}$, which returns 1 if the argument is equal to (t, a) and 0 otherwise. Thus, the dummy players' payoffs indicate the types and actions of the original, real players. In particular, for every correlated equilibrium distribution η and every element (t, a) of $T \times A$, the expected payoff of the corresponding dummy player is equal to $\eta(\{(t, a)\})$. It follows that two CEDs in the extended game, η and $\tilde{\eta}$, give the same CEP if and only if they are equal, $\eta = \tilde{\eta}$.

Every mechanism in the original game can be extended in a natural way to a mechanism in the extended game by sending arbitrary constant messages to the dummy players. The original and extended mechanisms have the exact same properties in $\{S, \tilde{S}, O, \tilde{O}, V, I\}$, and in the following they are identified. Using this identification, every correlated strategy in the original game can be extended in a natural way to a correlated strategy with the same mechanism in the extended game by assigning to each of the dummy players his single possible strategy. Observe that:

1. the original correlated strategy has the same distribution as the extended one (recall the above comment regarding the identification of profiles in the original and the extended games), and
2. one of them is a correlated equilibrium if and only if this is so for the other.

Moreover, every CED in the extended game can be obtained in the above manner from some CED in the original game. The former may be the distribution of a correlated strategy with a mechanism that sends variable messages to some dummy players. However, these messages are inconsequential (since a dummy player has only one possible action) and hence can be replaced by constant ones. Such replacement preserves each of the properties $S, \tilde{S}, O, \tilde{O}, V$ and I.

Suppose now that $\mathcal{P} \underset{\text{CEP}}{\Rightarrow} \mathcal{Q}$. Then, for every $\mathcal{P}$-implementable CED η in the extended game there is a $\mathcal{Q}$-implementable CED $\tilde{\eta}$ in the same game with the same payoff vector. As indicated, necessarily $\tilde{\eta} = \eta$, so that η is also $\mathcal{Q}$-implementable. It follows, by Observations 1 and 2 above, that every $\mathcal{P}$-implementable CED in the original game is also $\mathcal{Q}$-implementable. This proves that $\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}$.

Inspection of the above proof of (52) reveals that it applies virtually unchanged also to the more general version in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$. ■

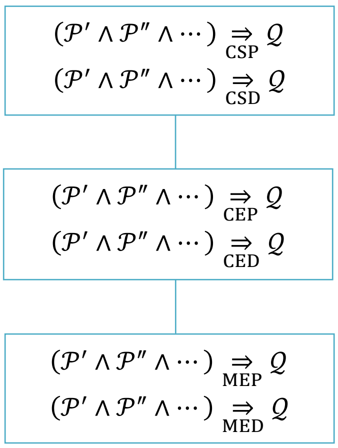

Figure 4. Hesse diagram of the implication relations between attributes of (equilibrium) outcomes in Bayesian games: correlated strategy payoffs (CSPs), correlated strategy distributions (CSDs), correlated equilibrium payoffs (CEPs), correlated equilibrium distributions (CEDs), communication equilibrium payoffs (MEPs) and communication equilibrium distributions (MEDs). A box represents a pair of equivalent implication relations: for all subsets $\mathcal{P}, \mathcal{P}', \mathcal{P}''$, … and $\mathcal{Q}$ of (the set of properties of mechanisms) $\{S, \tilde{S}, O, \tilde{O}, D, I\}$, one implication holds is and only if the other holds. A line represents only one-way implication (between implications): whenever an implication relation in the lower box holds, so does each implication relation in the higher box.

8 Summary

The three Hesse diagrams in Section 3 refer to different notions of outcomes or solution concepts. Specifically, they present the implication relation between attributes of correlated strategy distributions or payoffs (Figure 1), between attributes of correlated equilibrium distributions or payoffs (Figure 2), and between attributes of communication equilibrium distributions or payoffs (Figure 3). However, since in all six cases the attributes are defined in terms of properties of the implementing mechanisms, the implication relations themselves are potentially comparable. Indeed, the results in the previous sections (Propositions 11 and 22 and Theorem 4) show that each of these relations implies or is implied (or both) by each of the others. This implication relation between implication relations is shown by the Hesse diagram in Figure 4.

The Hesse diagrams in Figure 1, Figure 2, Figure 3 and Figure 4 together imply that the number of classes of CSDs (and CEPs) is strictly smaller than the number for CEDs (and CEPs), and the latter is strictly smaller than the number of classes of MEDs (and MEPs). Only the first of these numbers is precisely known: seven. The second number is either 14 or 15 (depending on the answer to the Open Problem presented in Section 5.3) and the third number (which, as indicated, is strictly greater that the second) is 15, 16 or 17.

Correlated and communication equilibria can both be viewed as special cases of a model in which the messages that the mechanism sends to the players may depend on both their true and reported types. The dependence on the latter may be of little significance if there are no limitations on the mechanism's use of the former. However, the present setup is constructed specifically for facilitating analyzing such limitations and their significance. Thus, suppose for example that only certain aggregate data regarding the players' true types are available to the mechanism, e.g., a "checksum" of the types. Then, unilateral deviations from truthful type reporting may be detectable even if the profile of reported type in not itself impossible, but not so for the identity of the player who lied about his type. The meaning of a correlated strategy in this general setting is the same as in the two special ones: it translates the mechanism's messages into type-dependant actions for the players. A natural requirement, which generalizes both correlated and communication equilibrium, is that it is incentive compatible for players to truthfully report their types and take the actions that are indicated by the messages they receive. The question arises, how different limitations on the mechanism affect the outcomes of such "correlated–communication" equilibria. From this perspective, the results reported in this paper only concern two special kinds of limitations. In the first, the reported types cannot affect the messages that the mechanism sends to the players, and in the second, the true types cannot affect the messages.

References

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55, 1–18.

Ben-Porath, E. (2003). Cheap talk in games with incomplete information. *Journal of Economic Theory* 108, 45–71.

Bergemann, D. and Morris, S. E. (2007). Belief free incomplete information games. Cowles Foundation Discussion Paper No. 1629, Yale University.

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

Cotter, K. D. (1994). Type correlated equilibria for games with payoff uncertainty. *Economic Theory* 4, 617–627.

Forges, F. (1990). Universal mechanisms. *Econometrica* 58, 1341–1364.

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

Forges, F. (2006). Correlated equilibrium in games with incomplete information revisited. *Theory and Decision* 61, 329–344.

Gerardi, D. (2004). Unmediated communication in games with complete and incomplete information. *Journal of Economic Theory* 114, 104–131.

Kar, A., Ray, I. and Serrano, R. (2010). A difficulty in implementing correlated equilibrium distributions. *Games and Economic Behavior* 69, 189–193.

Lehrer, E., Rosenberg, D. and Shmaya, E. (2006). Signaling and mediation in Bayesian games. Tel Aviv University. Mimeo.

Lehrer, E., Rosenberg, D. and Shmaya, E. (2010). Signaling and mediation in games with common interests. *Games and Economic Behavior* 68, 670–682.

Milchtaich, I. (2004). Random-player games. *Games and Economic Behavior* 47, 353–388.

Myerson, R.B. (1994). Communication, correlated equilibria, and incentive compatibility. Chapter 24 in *Handbook of Game Theory*, Vol. 2, edited by R. J. Aumann and S. Hart, Elsevier Science, pp. 827–847.

Samuelson, L. and Zhang, J. (1989). Correlated and mediated equilibria in games with incomplete information. Pennsylvania State University.

Shiryaev, A. N. (1996). *Probability*, 2nd Edition. Springer, New York.