Computationally Efficient Coordination in Game Trees

Françoise Forges

THEMA, Université de Cergy-Pontoise, and Institut Universitaire de France
email: Francoise.Forges@eco.u-cergy.fr

Bernhard von Stengel

Department of Mathematics, London School of Economics
email: stengel@maths.lse.ac.uk

April 25, 2002

CDAM Research Report LSE-CDAM-2002-02

Abstract. The solution concept of "correlated equilibrium" allows for coordination in games. For game trees with imperfect information, it gives rise to NP-hard problems, even for two-player games without chance moves. We introduce the "extensive form correlated equilibrium" (EFCE), which extends Aumann's correlated equilibrium, where coordination is achieved by signals that are received "locally" at information sets. An EFCE is polynomial-time computable for two-player games without chance moves.

1 Background

Game theory is the formal study of conflict and cooperation. It provides a language to formulate, structure and analyze scenarios where the actions of several agents are interdependent. Game-theoretic concepts are a major tool in theoretical economics [25]. In theoretical computer science, two-player games are familiar in models of complexity [6], for proving lower bounds for randomized algorithms [40], and in the competitive analysis of online algorithms [5]. Game theory is also recognized as a main tool for modeling and analyzing interactions on the internet. This is forcefully argued by Papadimitriou [29], who also notes significant computational questions posed by game theory. One of these open problems is the complexity of finding one Nash equilibrium of a two-player game given in strategic form.

The *strategic form* (also called normal form) is a basic model studied in noncooperative game theory. A game in strategic form is given by a set of strategies for each player, and specifies the payoff for each player resulting from each *strategy profile* (a strategy profile is a combination of strategies, one for each player). The game is played *simultaneously* by each player choosing a strategy, unaware of the choices of the other players, whereupon the players receive their payoffs.

The predominant solution concept for strategic-form games is the *Nash equilibrium* [27]. This is a strategy profile such that no player can improve his payoff by unilaterally changing his strategy. In order for Nash equilibria to exist, it may be necessary that players use *mixed strategies*. A mixed strategy of a player is given by a randomization over the given set of "pure" strategies of that player. A mixed strategy profile is a Nash equilibrium if no player can get a better *expected* payoff, assuming that the strategies of the other players stay fixed.

Any finite strategic-form game has a Nash equilibrium in mixed strategies [27]. The known constructive proofs, however, lead at best to exponential-time algorithms for finding one equilibrium. For two-player games, a classical algorithm that finds one Nash equilibrium, due to Lemke and Howson [22], relies on a parity argument [28] for the odd-degree nodes of a graph. This graph is defined by edges of polytopes [32, 38] and the algorithm is similar to the simplex method for linear programming, and has exponential worst-case running time. As of yet, no such combinatorial "pivoting" algorithms for linear programming are known that have polynomial worst-case running time. The more general problem of finding a single Nash equilibrium is therefore likely to be even more difficult and is considered as one of the most important concrete open questions on the boundary of P today [29].

The set of all Nash equilibria of a game is disconnected and computationally difficult in the sense that maximizing a linear function of the payoffs of the players is NP-hard [13] (this is independent from the mentioned question of finding *one* equilibrium). The concept of "correlated equilibrium", which generalizes Nash equilibrium, however, is computationally more tractable since the set of correlated equilibria of a game is a convex polytope.

A *correlated equilibrium*, due to Aumann [2], differs from a Nash equilibrium in that it allows for coordinated random choices of the players. A commonly known joint distribution on strategy profiles is used to select one of these profiles, whereupon each player is told only his strategy in that profile. The selection of the profile requires some device or *mediator*. After the players have learned their strategy, each has a posterior conditional distribution on what the other players have been recommended to do. Assuming they follow this recommendation, the equilibrium condition states that the player must have no incentive to deviate from the own recommended strategy. These *incentive constraints* can be described by linear inequalities, derived from the payoffs, with the joint probabilities for the strategy profiles as variables (see (1) below). They compare any two strategies of a player and are hence quadratic in the size of the game. The set of correlated equilibria is therefore a polyhedron defined by a polynomial number of linear inequalities. A correlated equilibrium with maximum payoff sum, for example, can therefore [14] be found in polynomial time [13].

A motivation for correlated equilibrium is that it describes the strategic possibilities of *pre-play communication* between the players [26]. This communication is "cheap talk" in that it does not involve any commitment, which would alter the game much more substantially. One can view correlated equilibrium as the Nash equilibrium of a game derived from the original game with an additional pre-play communication stage where players may exchange messages. The mediator can even be made unnecessary by using private pairwise interactions [3, 4], or suitable cryptographic protocols [9, 34]. As a further motivation, certain learning mechanisms converge to correlated equilibria [16].

This paper studies correlated equilibria for a game model that is much more detailed than the strategic form, the game tree or *extensive game*. This game representation is popular as it allows detailed and accurate models of strategic interactions. In an extensive game, a tree describes the game states (as nodes) and players' moves (as tree edges). Nodes may also belong to *chance* selecting the next node according to known probabilities. A game play starts at the root and ends at a leaf of the tree, where each player receives a payoff. *Imperfect information* in an extensive game is modeled by *information sets* [20]. An information set is a set of nodes that all have the same player to move and the same choices (denoted by labels on tree edges) at each of those nodes. A player is informed only about the information set she is at but not at which node, and her move is by definition the same at each of these nodes. This allows to model that a player is not fully informed about previous moves of other players. In a game of perfect information (like chess), all information sets are singletons and can be identified with the players' decision nodes.

A *strategy* in an extensive game is defined as a tuple of moves, one for each information set of the player. The strategic form of the game is obtained by listing the payoffs, or expected payoffs if there are chance moves, that result in the tree for any strategy profile. A Nash equilibrium is then defined as before.

Standard methods for finding Nash equilibria apply to the strategic form of the extensive game. If the game tree is the input, this is computationally very inefficient since the number of strategies is clearly *exponential* in the number of information sets of a player, and hence typically exponential in the size of the game tree. The *reduced strategic form* considers only "reduced" strategies omitting moves at information sets that are unreachable due to an earlier move of the player. The resulting growth is typically sub-exponential, like $2^{\sqrt{n}}$ for a typical tree with n nodes [39], so that using the reduced strategic form only becomes impractical for trees with several hundred nodes. For larger trees, the exponential "explosion" makes algorithms based on the reduced strategic form intractable, which has also limited the use of extensive games in practice [23].

A strategic description of *linear size* in the size of the game tree is the *sequence form* of an extensive game, due to [18, 36] and, in retrospect, [30]. It is based on sequences of moves, which are the moves of a particular player along a path in the game tree. The sequences are played according to certain "realization" probabilities, which are characterized by linear equations, one for each information set of a player (see equations (2) below). The resulting *realization plans* are the analog of mixed strategies for the sequence form. They can be translated to *behavior strategies* [20], which describe how to randomly choose moves at an information set. It is this "local" randomization of a behavior strategy that reduces the complexity from exponential to linear, as opposed to the "global", and very redundant, description by a mixed strategy that first picks one of the exponentially many pure strategies which is then used by the player in the tree.

Behavior strategies are due to Kuhn [20]. They are as powerful as mixed strategies if each player has *perfect recall*. This means that the information sets reflect that the player does not forget what he knew or did earlier, which is a standard assumption (the interpretation of game trees with imperfect recall becomes difficult, see e.g. [15]). The sequence form also requires perfect recall. It replaces behavior probabilities for moves by the probabilities for sequences ending in those moves. In two-person games, this gives rise to linear inequalities for the expected payoffs. Two-player zero-sum extensive games can then be solved in polynomial time. This was first stated explicitly in [17] who used sequences for one player and strategies with a separation oracle [14] for the other player; Romanovskii's earlier result [30] was still overlooked at that time. For two-player non-zero-sum games, Lemke's pivoting algorithm [21] can be used [19], an extension of the Lemke–Howson method for strategic-form games [22]. A (more general) implementation with exact integer arithmetic performs well in computational experiments [39]. The sequence form has been used for solving large game trees with linear programming to obtain new lower bounds in the list update problem [1], a classical online problem [5].

2 Contribution of this paper

This paper investigates *polynomial-time computable* correlated equilibria for extensive games with perfect recall. As described above, polynomial algorithms for strategic-form games are known for finding Nash equilibria of *zero-sum* two-person games, and for finding *correlated* equilibria. For non-zero-sum games, the complexity of finding one Nash equilibrium is open (with a pivoting algorithm [22] that works well in practice), and problems relating to the set of all Nash equilibria tend to be NP-hard [13]. For two-player extensive games with perfect recall, the sequence form makes Nash equilibria computationally as tractable as for games in strategic form [30, 18, 36, 39].

Is there a "sequence form" to compute correlated equilibria of extensive games efficiently? The answer is *negative* when considering two-player extensive games with perfect recall and chance moves: Chu and Halpern [7] recently established that finding a maximum-payoff-sum correlated equilibrium for such games is NP-hard to compute, even if the players have identical payoffs. The set of correlated equilibria can therefore not be characterized by a polynomial number of inequalities in the size of the game tree, unless P $=$ NP. The proof of this result converts a SAT instance to an extensive game ([7] actually uses a "possible worlds" model) where the strategic form is similar to a truth table for the SAT formula, with a chance move picking one of the clauses. The chance move can be replaced by an active randomization of one of the players, using an initial generalized "rock–scissors–paper" game, to yield an even stronger result:

Theorem 2.1 [37]**.** *For two-player, perfect-recall extensive games without chance moves, it is NP-hard to find a strategic-form correlated equilibrium with maximum payoff sum.*

The exponential number of pure strategies in an extensive game seems to be unavoidable when considering correlated equilibria, as long as these are defined in terms of the strategic form. Our main contribution is an *alternative definition* of correlated equilibrium for extensive games, which we call *extensive form correlated equilibrium* or EFCE. It is similar to the known strategic-form correlated equilibrium in that it generates recommendations of moves *before* the game starts. However, a player receives the signal with the recommended move *when reaching an information set*, as if in a "sealed envelope" that she can open then, but not earlier.

The EFCE generalizes Nash equilibria in behavior strategies, and is closer in spirit to the dynamic description of the game by a tree than the strategic-form correlated equilibrium. At the same time, the game is altered minimally since the mediator generates the signals at the beginning of the game. Other extensions of correlated equilibrium have been proposed for *specific* classes of games, like Bayesian games [11, 31, 8, 12] or multi-stage games [10, 24]. In contrast, our concept seems to be the first that applies to *general* extensive games. For instance, "autonomous" correlated equilibria [10, 33] and "communication equilibria" [10, 24] are only defined for multistage games, as they rely on devices which give private recommendations to each player at every stage the game. In the case of communication equilibria, the players can send messages to the device at every stage. Even more general communication equilibria are considered in [33] where the device can also base recommendations on past play.

Any strategic-form correlated equilibrium is an EFCE, but the set of EFCE is in general larger. This is known in special cases [24, Fig. 2] and unsurprising since in an EFCE the players have less information and so incentives can be more easily met. In multistage games, any autonomous correlated equilibrium is an EFCE. However, the converse is not true unless further assumptions are made on the players' information [33]. It is easy to see that there is no inclusive relationship between communication equilibria and EFCE.

The definition of an EFCE applies to any extensive game. For two players and *no chance moves*, the set of EFCE *can be described by a polynomial number of linear constraints*, in the size of the game tree. This positive result (Theorem 5.1 below) is the most substantive feature of the EFCE concept from a computational viewpoint. Interestingly, EFCE are still hard to characterize when a chance player or a third player are allowed in the game:

Theorem 2.2 [7, 37]**.** *Consider a set of "equilibria" which are convex combinations of pure strategy pairs and include all Nash equilibria. Then for an extensive game with two players and chance moves, or with three players, it is NP-hard to find an "equilibrium" with maximum payoff sum.*

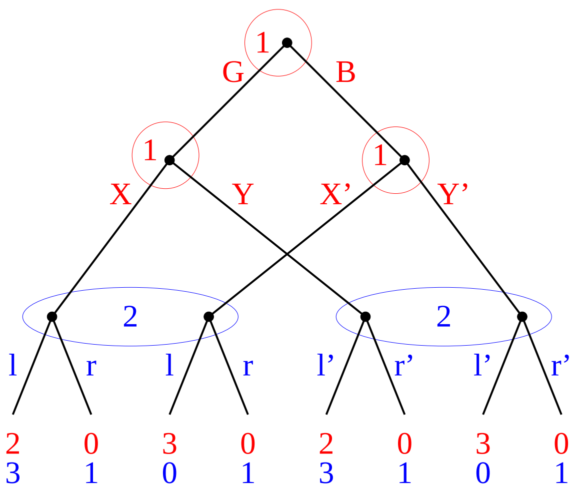
We will illustrate the EFCE concept in Section 3. Our main result is its polynomial-time computatibility for two-player games without chance moves. Section 4 covers the *consistency constraints* for generating moves that may be correlated across the different information sets of the two players. This is the main difficulty, since one has to avoid correlations of own moves, which would re-introduce the complex mixed strategies that cannot be described by a small number inequalities. Here, the absence of a third player or of chance, and the condition of perfect recall, is crucial. From the "correlated move probabilities" one can generate "locally" a pair of pure strategies, analogous to a behavior strategy of one player in a Nash equilibrium. The resulting distribution on pure strategy pairs forms an EFCE if it fulfills additional *incentive constraints*, described in Section 5. These compare the own expected payoff when following a recommended move (assuming the player will observe further recommendations in the future) with the possible payoff when *deviating* from the recommended move, which must not be higher. The deviation payoff is computed by optimizing in the tree against the behavior of the other player, according to the current conditional probabilities. This is akin to dynamic programming and captured by suitable variables associated with moves and information sets. Section 6 concludes.

3 Example of an extensive-form correlated equilibrium

Figure 1 shows an example of an extensive game. Player 1, a student, chooses a good (G) or bad (B) education, which defines his "type". Afterwards, he applies for a summer research job with a professor, player 2. Player 1 sends a signal X or Y (we add primes as in X' and Y' only to make choices at different information sets distinct). The professor can distinguish the signals but not the type of player 1, as shown by her two information sets. She can either let the student work with her (l) or refuse to do so (r). Move r always gives payoffs $(0, 1)$ to players 1 and 2, but l results in $(2, 3)$ for G versus $(3, 0)$ for B.

In games of incomplete information, the type is normally chosen by a chance move, not the player himself. However, larger games of this sort are not easy to solve in general, so that this game without chance moves demonstrates better our EFCE concept.

FIGURE 1



The Nash equilibria of this game are given as follows. Player 2 refuses to work with the student, with the strategy (r, r'), since any positive probability for l or l' would induce player 1 to choose B along with the appropriate signal X or Y, which is better than G. Then l or l' is certainly not optimal for player 2. Hence the choice of B or G and of the signal for player 1 do not matter (he gets payoff 0 anyhow), as long as in no information set of player 2, the probability for G versus B is high enough to make her switch to l.

This "economically inefficient" outcome of the game could be avoided if player 1 could choose G and signal this appropriately, without being able to mimic this when he is of type B. This requires coordination between the two players, as offered by a correlated equilibrium. However, it is not possible with any such concept based on the strategic form, or multiple stages [10, 24], where player 1 gets the recommendations for both types G and B. An EFCE, however, gives this possibility: Suppose the reduced pure strategy profiles $((G, X, *), (l, r'))$ and $((G, Y, *), (r, l'))$ are chosen with probability 1/2 each. The moves in these profiles are revealed to the players when reaching their respective information sets. Player 1 is not recommended to play B and hence gets no signal X' or Y', indicated by "$*$". After G, he knows that he will get a signal X or Y that is perfectly correlated with player 2's choice l or l' to let him work with her, giving him payoff 2. When deviating and choosing B, however, the signal will be not revealed, and X' and Y' will both have probability 1/2 for the response r or r', giving the expected payoff 3/2 which is less than 2 when following the recommendation, so player 1 indeed follows it. Player 2 gets recommendation l or l' and knows that player 1 is of the good type G when following his recommendation, so l and l' are also optimal for player 2.

4 Consistency constraints

Throughout, we consider an extensive two-person game with perfect recall and no chance moves. We will show that the set of EFCE for such a game can be described by a *small* number (polynomial in the size of the game tree) of linear constraints. The linear constraints will be *consistency constraints* that describe the possible probability distributions on profiles of moves to be recommended to the players, and additional *incentive constraints*, described in the next section, that assert when it is optimal for the players to follow these recommendations. As a prerequisite, we first review correlated equilibria for a two-player game in strategic form, and subsequently the sequence form of an extensive game as used for finding Nash equilibria.

A correlated equilibrium of a strategic-form two-player game can be defined as follows [2, 26]. Let i and j stand for strategies of player 1 and 2, respectively, with resulting payoffs a_{ij} and b_{ij}. A correlated equilibrium is a distribution on strategy pairs. When a strategy pair (i, j) is drawn according to this distribution, player 1 is told i and player 2 is told j. The probabilities Z_{ij} are nonnegative and sum up to one, which defines the *consistency constraints*. Furthermore, for all strategies i and k of player 1 and all strategies j and l of player 2,

$$\sum_j Z_{ij}\, a_{ij} \ge \sum_j Z_{ij}\, a_{kj}\,, \qquad \sum_i Z_{ij}\, b_{ij} \ge \sum_i Z_{ij}\, b_{il}\,. \tag{1}$$

The *incentive constraints* (1) state that player 1, when recommended to play i, has no incentive to switch from i to k, given (up to normalization) the conditional probabilities Z_{ij} on opponent strategies j. Analogously, the second inequalities in (1) state that player 2, when recommended to play j, has no incentive to switch to l.

The strategic-form description of an EFCE is computationally disadvantageous because the number of pure strategies is exponential in the size of the game tree. For finding Nash equilibria, the sequence form is of linear size. However, its randomized strategies, called "realization plans", are more complicated to describe than mixed strategies. Similarly, our characterization of EFCE with sequences will require more complicated consistency constraints than the strategic form.

We use a standard notation for extensive games [39]. The non-terminal *decision* nodes of the game tree are partitioned into *information sets*. Each information set belongs to exactly one player i. The set of all information sets of player i is denoted H_i. The set of choices or moves at an information set h is denoted C_h. Each node in h has $|C_h|$ outgoing edges, which are labeled with the moves in C_h. Choice sets C_h and C_k for $h \neq k$ are disjoint. The *sequence form* uses sequences of moves of a particular player as encountered along the path from the root to any node in the game tree. By definition, player i has *perfect recall* if all nodes in an information set h in H_i define the same sequence σ_h of moves for player i. Hence, any move c at h is the last move of a unique sequence $\sigma_h c$. This defines all possible sequences of a player except for the empty sequence $\emptyset$. The set of sequences of player i is denoted S_i, so

$$S_i = \{\, \emptyset \,\} \,\cup\, \{\, \sigma_h c \mid h \in H_i,\ c \in C_h \,\}.$$

For brevity, we also denote sequences of player 1 by σ and sequences of player 2 by τ, and the sequence leading to an information set h of player 2 by τ_h.

The sequence form is applied to Nash equilibria as follows [18, 36, 39]. Sequences are played randomly according to *realization plans*. A realization plan x for player 1 is given by nonnegative real numbers $x(\sigma)$ for $\sigma \in S_1$, a realization plan y for player 2 by nonnegative numbers $y(\tau)$ for $\tau \in S_2$. These denote the realization probabilities for the sequences σ and τ when the players use certain mixed strategies. For player 1, such a realization plan is characterized by the equations

$$x(\emptyset) = 1, \qquad \sum_{c \in C_h} x(\sigma_h c) = x(\sigma_h) \qquad (h \in H_1)\,, \tag{2}$$

and analogously for player 2 with y and H_2 instead of x and H_1. Equations (2) hold naturally when player 1 uses a behavior strategy, in particular a pure strategy, and hence also for a mixed strategy which is a convex combination of pure strategies. A realization plan x fulfilling (2) results from a behavior strategy that chooses move c at an information set $h \in H_1$ with probability $x(\sigma_h c)/x(\sigma_h)$ if $x(\sigma_h) > 0$ and arbitrarily if $x(\sigma_h) = 0$. This yields a canonical proof of Kuhn's theorem [20] that asserts that a player with perfect recall can replace any mixed strategy by an equivalent behavior strategy. The behavior at h is unspecified if $x(\sigma_h) = 0$, which means that h is unreachable due to an earlier own move. Not specifying the behavior at such information sets is exactly what is done in the reduced strategic form.

Because the game has no chance moves, any leaf of the game tree defines a unique pair (σ, τ) of sequences of the two players. Let $a(\sigma, \tau)$ and $b(\sigma, \tau)$ denote the respective payoffs to the players at that leaf. Then if the two players use the realization plans x and y, their expected payoffs are given by the bilinear expressions

$$\sum_{\sigma,\tau} x(\sigma)\, y(\tau)\, a(\sigma,\tau)\,, \qquad \sum_{\sigma,\tau} x(\sigma)\, y(\tau)\, b(\sigma,\tau)\,, \tag{3}$$

respectively. The expressions in (3) represent the sum, over all leaves, of the payoffs, multiplied by the probabilities of reaching the leaves. The sums in (3) may be taken over all $\sigma \in S_1$ and $\tau \in S_2$ by assuming that $a(\sigma, \tau) = b(\sigma, \tau) = 0$ whenever the sequence pair (σ, τ) does not lead to a leaf. This is useful when using matrix notation, where the payoffs in the sequence form are entries $a(\sigma, \tau)$ and $b(\sigma, \tau)$ of sparse $|S_1| \times |S_2|$ payoff matrices and x and y are regarded as vectors. Using linear programming duality, conditions for Nash equilibria can then be written in terms of payoffs and transposed constraints (2) which require one equation and one dual variable for each information set [18, 36]. This results into a small linear program for zero-sum payoffs, and a small linear complementarity problem for non-zero-sum payoffs [19].

In order to describe an EFCE, the product $x(\sigma)\, y(\tau)$ in (3) of the realization probabilities for σ in S_1 and τ in S_2 will be replaced by a more general *joint* realization probability $z(\sigma, \tau)$ that the pair of sequences (σ, τ) will be recommended to the two players, as far as this probability is relevant. These probabilities $z(\sigma, \tau)$ define what we call a *correlation plan* for the game.

In an EFCE, a player gets a move recommendation when reaching an information set. The move corresponds uniquely to a sequence ending in that move. For player 1, say, the sequence denotes a row of the $|S_1| \times |S_2|$ correlation plan matrix. From this row, player 1 should have a posterior distribution on the recommendations to player 2. This behavior of player 2 must be specified not only when player 1 follows a recommendation, but also when player 1 deviates, so that player 1 can decide if the own recommendation is optimal. The recommendations to player 2 off the equilibrium path are therefore important. Otherwise, one could simply choose a distribution on the leaves of the tree (with a correlation plan that is sparse like the payoff matrix), and merely recommend to the players the pair of sequences corresponding to the selected leaf. This does not suffice, since an EFCE must recommend *strategies* to the players.

Our first approach is therefore to define a correlation plan z as a full matrix. Up to normalization (which is not needed in (1) either), a column of this matrix should be a realization plan of player 1 and a row a realization plan of player 2. Omitting the normalizing first equation in (2), this means that for all $\tau \in S_2$, $h \in H_1$, $\sigma \in S_1$, and $k \in H_2$,

$$\sum_{c \in C_h} z(\sigma_h c, \tau) = z(\sigma_h, \tau), \qquad \sum_{d \in C_k} z(\sigma, \tau_k d) = z(\sigma, \tau_k). \tag{4}$$

Furthermore, the pair $(\emptyset, \emptyset)$ of empty sequences is selected with certainty, and the probabilities are nonnegative, which adds the trivial consistency constraints

$$z(\emptyset, \emptyset) = 1, \qquad z(\sigma, \tau) \geq 0 \qquad (\sigma \in S_1, \tau \in S_2). \tag{5}$$

The constraints (4) and (5) hold for the special case $z(\sigma, \tau) = x(\sigma)y(\tau)$ where x and y are realization plans. With properly defined incentive constraints, such a correlation plan of rank one should define a Nash equilibrium, just as a strategic-form correlated equilibrium with a rank-one matrix Z in (1) is a Nash equilibrium. In particular, if x and y stand for pure strategies, where each sequence σ or τ is chosen with probability zero or one, then the probabilities $z(\sigma, \tau) = x(\sigma)y(\tau)$ are also zero or one. For any *convex combination* of pure strategies, as in an EFCE, (4) and (5) therefore hold as well, so these are *necessary* conditions for a correlation plan.

Figure 2 shows a correlation plan arising from a pure strategy pair, for the game in Figure 1 when the first move of player 1 is replaced by a chance move. Figure 3 shows a possible assignment of probabilities $z(\sigma, \tau)$ that fulfills (4) and (5). These probabilities are "locally consistent" in the sense that the marginal probability of each move is 1/2. However, they *cannot* be obtained as a convex combination of pure strategy pairs as in Figure 2. Otherwise, one such pair would have to recommend move X to player 1 and move l to player 2 to account for the respective entry 1/2. In that pure strategy pair, given that player 2 is recommended move l, the recommendation to player 1 at the other information set must be Y' since the move combination (X', l) has probability

	$\emptyset$	l	r	l'	r'
$\emptyset$	1	1	0	1	0
X	1	1	0	1	0
Y	0	0	0	0	0
X'	0	0	0	0	0
Y'	1	1	0	1	0

FIGURE 2

	$\emptyset$	l	r	l'	r'
$\emptyset$	1	1/2	1/2	1/2	1/2
X	1/2	1/2	0	1/2	0
Y	1/2	0	1/2	0	1/2
X'	1/2	0	1/2	1/2	0
Y'	1/2	1/2	0	0	1/2

FIGURE 3

zero. Similarly, move X requires that move l' is recommended to player 2. This pure strategy pair is thus $((X, Y'), (l, l'))$ as in Figure 2, but that also selects (Y', l'), contradicting Figure 3. This shows that (4) and (5) do not suffice to characterize the convex hull of pure strategy profiles. For a game with chance moves, the NP-hardness in Theorem 2.2 shows that this convex set cannot be characterized by a polynomial number of linear inequalities, unless P $=$ NP.

For a two-player game without chance moves, however, this problem can be resolved by specifying only correlations of moves at "connected" information sets where decisions can affect each other during play. Call any two information sets h and k (even of the same player) *connected* if there is a path from the root to a leaf containing a node of h and a node of k. If the node in h comes earlier on the path, then h is said to *precede* k. The following lemma states that the two-player games without chance moves considered here have a weak "temporal" structure in the sense that a player can always tell if he is to move before or after the other player.

Lemma 4.1. *For any two information sets h and k, if h precedes k, then k does not precede h.*

Amending our first approach, we define a *correlation plan* $z\colon S_1 \times S_2 \to \mathbb{R}$ as follows. First, there is a joint probability distribution on the set of reduced pure strategy pairs (π_1, π_2) of the two players so that $z(\sigma, \tau)$ is the combined probability of the strategy pairs (π_1, π_2) where π_1 agrees with σ (that is, chooses all the moves in σ) and π_2 agrees with τ. Second, z is a *partial* function where $z(\sigma, \tau)$ is specified only for "relevant" sequence pairs (σ, τ). The pair (σ, τ) in $S_1 \times S_2$ is called *relevant* if σ or τ is the empty sequence, or if $\sigma = \sigma_h c$ and $\tau = \tau_k d$ for connected information sets h and k, where $h \in H_1$, $c \in C_h$, $k \in H_2$, $d \in C_k$. Note that the *information sets* are connected where the respective last move in σ and τ is made. It is not necessary that the sequences themselves share a path. We specify correlations of moves at connected information sets, not just of moves sharing a path, since a player may consider deviations from the recommended moves. The following shows that equations (4) can be sensibly restricted to relevant sequence pairs.

Lemma 4.2. *If the pair (σ, τ) of sequences is relevant, and σ' is a prefix of σ and τ' is a prefix of τ, then (σ', τ') is relevant.*

In this way, we obtain the consistency constraints for correlation plans. The correlation plan itself can also be used to generate, as a random variable, a pair of strategies to be recommended to the two players.

Theorem 4.3. *In a two-player, perfect-recall extensive game without chance moves, z is a correlation plan if and only if it fulfills* (5)*, and* (4) *whenever $(\sigma_h c, \tau)$ and $(\sigma, \tau_k d)$ are relevant, for any $c \in C_h$ and $d \in C_k$. A corresponding joint probability distribution on pairs of reduced pure strategies can be generated directly from z.*

5 Incentive constraints

In an EFCE, a player gets a move recommendation when reaching an information set. This recommendation induces a posterior distribution on the recommendations given to the other player. For past moves, this induces a certain distribution on where the player is in the information set. For future moves, it expresses the subsequently expected play. Both are represented by the eventual distribution on the leaves of the game tree. The players want to optimize the expected payoffs which they receive at the leaves, assuming the other player follows her recommendations.

The *incentive constraints* in an EFCE express that it is optimal to follow any move recommendation, under two assumptions about the player's *own* behavior: When *following* the recommended move, the player considers the *expected* payoff when following recommendations in the future. When *deviating* from the recommended move, the player *optimizes* his payoff, given the current knowledge about the other player's behavior. Any recommendations given *after* a deviation are ignored, and are in fact not given, since an EFCE only generates a pair of *reduced strategies*: When a player deviates, he subsequently only reaches own information sets that would be unreachable when following the original move in the strategy, so these later moves are left unspecified in a reduced strategy.

The sequence form only allows specifications of reduced strategies. Assume that a pair of reduced strategies is generated according to a correlation plan as in Theorem 4.3. Suppose that player 1, say, gets a recommendation for move c at an information set h, corresponding to the sequence $\sigma = \sigma_h c$. For the sequences τ of player 2, the row entries $z(\sigma, \tau)$ of the correlation plan z define, up to normalization, a realization plan that describes player 2's conditional behavior. This is only given for information sets connected to h, where (σ, τ) is relevant, which suffices for any decision of player 1 at this point.

We first introduce auxiliary variables $u(\sigma)$ for any $\sigma \in S_1$ (and, throughout, analogously for player 2). These denote the expected payoff contribution of σ (that is, of all strategies agreeing with σ) when player 1 follows recommendations. They are given by

$$u(\sigma) = \sum_{\tau} z(\sigma, \tau)\, a(\sigma, \tau) + \sum_{k \in H_1 :\, \sigma_k = \sigma} \; \sum_{d \in C_k} u(\sigma_k d)\,. \tag{6}$$

In (6), $a(\sigma, \tau)$ is the payoff to player 1 at the leaf defining the sequence pair (σ, τ), which is then obviously a relevant pair; if there is no such leaf, $a(\sigma, \tau) = 0$. The first sum in (6) captures the expected payoff contribution where σ and suitable sequences τ of player 2 are defined by leaves. The second, double sum in (6) concerns the information sets k of player 1 reached by σ. The sum of the payoff contributions $u(\sigma_k d)$ for $d \in C_k$ is the expected payoff when player 1 follows the recommendation to choose d at k, given the new posterior information obtained there.

Applying (6) recursively, starting with the longest sequences, gives for the empty sequence $u(\emptyset) = \sum_{\sigma,\tau} z(\sigma, \tau)\, a(\sigma, \tau)$. This denotes the overall payoff for player 1 under the correlation plan z (and similarly for player 2), which generalizes (3).

The payoff $u(\sigma)$ when following the recommended move c in $\sigma = \sigma_h c$ must be compared with the possible payoff when deviating from c. This is described by an optimization against the behavior of player 2 in row $\sigma_h c$ of z, by considering the other moves at h, as well as moves at information sets k that are reached later on. By optimizing in this way, the payoff contribution at an information set k of player 1 is denoted by $v(k, \sigma_h c)$. The parameter $\sigma_h c$ indicates the given row of the correlation plan z against which player 1 optimizes. For $k = h$, we define

$$v(h, \sigma_h c) = u(\sigma_h c). \tag{7}$$

The recommended move c should be optimal at h. This incentive constraint is expressed by the following inequalities, for any information set k in H_1 with $k = h$ or h preceding k, and all moves d at k:

$$v(k, \sigma_h c) \geq \sum_{\tau} z(\sigma_h c, \tau)\, a(\sigma_k d, \tau) + \sum_{l \in H_1 \colon \sigma_l = \sigma_k d} v(l, \sigma_h c) \qquad (d \in C_k). \tag{8}$$

The first sum in (8) is well defined, since when $(\sigma_k d, \tau)$ leads to a leaf, then $(\sigma_h c, \tau)$ is relevant because σ_h is a prefix of $\sigma_k d$. If there are no further information sets l and $k = h$, then (8) is analogous to (1), with moves c, d instead of strategies i, k. Here as there, the posterior distribution from the given recommendation $\sigma_h c$ is used for the comparison with other choices. In general, a move d at k can lead to further information sets l, also preceded by h, where the best possible payoff contribution is computed as $v(l, \sigma_h c)$. This variable is based on the *same* behavior of player 2 given by row $\sigma_h c$ of z.

The number of variables $v(k, \sigma_h c)$ is quadratic in the number of sequences of player 1 because they are indexed by the information sets k and the sequences $\sigma_h c$. The latter reflect the conditional behavior of the other player, which varies in a correlated equilibrium. In a Nash equilibrium, it would be fixed, and $z(\sigma_h c, \tau)$ is replaced by $y(\tau)$ for an unconditional realization plan y of player 2. Furthermore, the variables $v(k, \sigma_h c)$ are replaced by single variables $v(k)$, one for each information set k of player 1. Then the inequalities (8) are exactly those expressing the Nash equilibrium condition, with *dual* variables $v(k)$. These dual variables also express, like here, the optimization by "dynamic programming" [36, p. 239].

Together with the consistency constraints, the incentive constraints above characterize an EFCE. We summarize our main result as follows.

Theorem 5.1. *In a two-player, perfect-recall extensive game without chance moves, a correlation plan z as in Theorem 4.3 that fulfills the incentive constraints (6), (7), (8) defines an EFCE. The number of variables and constraints is polynomial in the size of the game tree, so that an EFCE is polynomial-time computable.*

6 Conclusions

A substantive game-theoretic setup seems required even to describe the problem addressed in this paper. However, the basic framework of correlated equilibria, namely coordination, communication, and incentives, is pervasive in economic theory, in particular mechanism design [25]. Our results are also of interest for pure game theory, where adapting equilibrium concepts to the dynamic tree structure has a long history, both for Nash [35] and correlated equilibria [10, 24, 12, 33].

Most saliently, our approach profits from, and contributes to, the *interaction* of computer science and game theory. Computational tractability of game-theoretic concepts is increasingly appreciated in economics [39]. The new EFCE concept is not only "natural" but also useful in being *easy to compute* for two-player game trees without chance moves. It seems new that, unlike, for example, in zero-sum games [18], the absence of chance can mark the borderline of P.

References

[1] Ambühl, C., B. Gärtner, und B. von Stengel (2001), A new lower bound for the list update problem in the partial cost model. *Theoretical Computer Science* 268, 3–16.

[2] Aumann, R. J. (1974), Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics* 1, 67–96.

[3] Bárány, I. (1987), Fair distribution protocols or how the players replace fortune. *Mathematics of Operations Research* 17, 327–340.

[4] Ben-Porath, E. (1998), Correlation without mediation: expanding the set of equilibrium outcomes by "cheap" pre-play procedures. *Journal of Economic Theory* 80, 108–122.

[5] Borodin, A. and R. El-Yaniv (1998), *Online Computation and Competitive Analysis*. Cambridge Univ. Press, Cambridge.

[6] Chandra, A. K., D. C. Kozen, and L. J. Stockmeyer (1981), Alternation. *J. ACM* 28, 114–133.

[7] Chu, F. and J. Halpern (2001), On the NP-completeness of finding an optimal strategy in games with common payoffs. *International Journal of Game Theory* 30, 99–106.

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

[9] Dodis, Y., S. Halevi, and T. Rabin (2000), A cryptographic solution to a game theoretic problem. In: *Proc. Crypto 2000*, Springer LNCS 1880, 112–130.

[10] Forges, F. (1986), An approach to communication equilibria. *Econometrica* 54, 1375–1385.

[11] Forges, F. (1986), Correlated equilibria in repeated games with lack of information on one side: a model with verifiable types. *International Journal of Game Theory* 15, 65–82.

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

[13] Gilboa, I. and E. Zemel (1989), Nash and correlated equilibria: some complexity considerations. *Games and Economic Behavior* 1, 80–93.

[14] M. Grötschel, L. Lovász, and A. Schrijver (1993), *Geometric Algorithms and Combinatorial Optimization*, 2nd ed. Springer-Verlag, Berlin.

[15] Halpern, J. Y. (2001), A computer scientist looks at game theory. Mimeo, Cornell University.

[16] Hart, S., and A. Mas-Colell (2000), A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68, 1127–1150.

[17] Koller, D. and N. Megiddo (1992), The complexity of two-person zero-sum games in extensive form. *Games and Economic Behavior* 4, 528–552.

[18] Koller, D., N. Megiddo, and B. von Stengel (1994), Fast algorithms for finding randomized strategies in game trees. In: *Proc. 26th STOC*, 750–759.

[19] Koller, D., N. Megiddo, and B. von Stengel (1996), Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior* 14, 247–259.

[20] Kuhn, H. W. (1953), Extensive games and the problem of information. In: *Contributions to the Theory of Games II*, eds. H. W. Kuhn and A. W. Tucker, Annals of Mathematics Studies 28, Princeton Univ. Press, Princeton, 193–216.

[21] Lemke, C. E. (1965), Bimatrix equilibrium points and mathematical programming. *Management Science* 11, 681–689.

[22] Lemke, C. E. and J. T. Howson, Jr. (1964), Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics* 12, 413–423.

[23] Lucas, W. F. (1972), An overview of the mathematical theory of games. *Management Science* 18, Appendix P, 3–19.

[24] Myerson, R. B. (1986), Multistage games with communication. *Econometrica* 54, 323–358.

[25] Myerson, R. B. (1991), *Game Theory – Analysis of Conflict*. Harvard Univ. Press, Cambridge, Mass.

[26] Myerson, R. B. (1994), Communication, correlated equilibria and incentive compatibility. In: *Handbook of Game Theory, Vol. 2*, eds. R. J. Aumann and S. Hart, Elsevier, Amsterdam, 827–847.

[27] Nash, J. F. (1951), Non-cooperative games. *Annals of Mathematics* 54, 286–295.

[28] Papadimitriou, C. H. (1994), On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Sciences* 48, 498–532.

[29] Papadimitriou, C. H. (2001), Algorithms, games, and the internet. In: *Proc. 33rd STOC*, 749–753.

[30] Romanovskii, I. V. (1962), Reduction of a game with complete memory to a matrix game. *Doklady Akademii Nauk* SSSR 144, 62–64 [English translation: *Soviet Mathematics* 3, 678–681].

[31] Samuelson, L., and J. Zhang (1989), Correlated equilibria and mediated equilibria in games with incomplete information. Mimeo, Penn State University.

[32] Shapley, L. S. (1974), A note on the Lemke–Howson algorithm. *Mathematical Programming Study 1: Pivoting and Extensions*, 175–189.

[33] Solan, E. (2001), Characterization of correlated equilibria in stochastic games. *International Journal of Game Theory* 30, 259–277.

[34] Urbano, A., and J. E. Vila (2002), Computational complexity and communication: coordination in two-player games. *Econometrica*, to appear.

[35] van Damme, E. (1987), *Stability and Perfection of Nash Equilibria*. Springer, Berlin.

[36] von Stengel, B. (1996), Efficient computation of behavior strategies. *Games and Economic Behavior* 14, 220–246.

[37] von Stengel, B. (2001), Computational complexity of correlated equilibria for extensive games. *CDAM Research Report LSE-CDAM-2001-03*.

[38] von Stengel, B. (2002), Computing equilibria for two-person games. Chapter 45, *Handbook of Game Theory, Vol. 3*, eds. R. J. Aumann and S. Hart, North-Holland, Amsterdam, in press.

[39] von Stengel, B., A. H. van den Elzen, and A. J. J. Talman (2002), Computing normal form perfect equilibria for extensive two-person games. *Econometrica* 70, 693–715.

[40] Yao, A. C. (1977), Probabilistic computation: Towards a unified measure of complexity. In *Proc. 18th FOCS*, 222–227, 1977.

Appendix: Proof outlines

In this appendix, we sketch the proofs of Theorems 4.3 and 5.1. We emphasize the conceptual aspects, in particular the crucial assumption of two players with perfect recall and no chance moves for Theorem 4.3.

Proof sketch of Theorem 4.3. In an EFCE, a move is recommended at each information set, where that recommendation is revealed to the player. Given a correlation plan z according to the equations (4) and (5) for relevant sequence pairs, we define inductively a move probability for each information set, assuming that moves have been recommended for all *preceding* information sets. The induction therefore proceeds "top down" from the root towards the leaves of the tree. It cannot "get stuck" because of Lemma 4.1.

Consider an information set h, say of player 1, under the inductive assumption that move recommendations have been generated for all information sets preceding h. If not all moves in the unique sequence σ_h leading to h have been recommended, then h will not be reached when player 1 follows the recommendations. In that case, a reduced strategy for player 1 will leave the move at h unspecified, and consequently no move will be recommended for h.

So suppose that all moves in σ_h have been generated as recommendations so far. The move recommendation at h has to be made in agreement with the recommendations given to player 2. Since these have been generated for all preceding information sets of player 2, we will now identify a *unique* such information set k in H_2. The moves at h will then be correlated with k, that is, they will be generated according to the joint probabilities, as specified in the correlation plan z, of sequence pairs ending in $C_h \times C_k$. (If the game had chance moves, it would not be possible to identify such a unique information set k of the other player, as Figure 2 shows, which refers to Figure 1 with the first move of player 1 replaced by chance. A correlation plan for the given game in Figure 1 would look different, with sequences $\emptyset, G, B, GX, GY, BX', BY'$ for player 1.)

The nodes in h (if there are several) define different sequences for player 2, since they define the same sequence σ_h for player 1 and no other player or chance is involved. The union of all the paths leading to the nodes in h defines a subtree T_h of the game tree. All branching points (nodes with more than one child) of T_h belong to player 2, since player 1 has perfect recall. These decision points belong to information sets of player 2 where move recommendations have already been given. Following the move recommendations from the root to h, we grow a sequence τ for player 2 until the generated path leaves the tree T_h, if ever. That is, the last move in τ is made at a node in T_h, and hence at an information set k preceding h. The last move in τ itself need no longer be on the path to h. If no information sets of player 2 precede h, then h is a singleton and we set $\tau = \emptyset$. The recommended move c at h is then determined according to the probability

$$\frac{z(\sigma_h c, \tau)}{z(\sigma_h, \tau)} \qquad (c \in C_h). \tag{9}$$

By construction of τ, the sequence pair $(\sigma_h c, \tau)$ in (9) is relevant, and the denominator is positive since both σ_h and τ have been recommended. (This assumes, inductively, that the numbers $z(\sigma, \tau)$ are indeed the probabilities that these sequence pairs $(\sigma_h c, \tau)$ will be recommended to the players.) Moreover, (9) defines a probability distribution on C_h by (4) and (5).

By looking at all information sets in this way and generating moves at each information set that is reachable due to the own earlier moves, a pair (π_1, π_2) of reduced pure strategies is generated, which is a random variable. It remains to show that this does not depend on the order in which one looks at the information sets, and that for *any* pair (σ, τ) of relevant sequences, $z(\sigma, \tau)$ is the probability that it will be recommended. The main observation here is that if either of two information sets can be chosen in the inductive construction, then the moves at these two sets are generated *independently*. This generalizes the independence observed in behavior strategies as used for Nash equilibria. □

Proof sketch of Theorem 5.1. The main arguments for the correctness of the incentive constraints (6), (7), (8) have already been given in Section 5. In particular, when a player deviates, then the assumed behavior of the other player is based on the last own recommendation that the player received, in agreement with the preceding proof sketch of Theorem 4.3.

We mention here only an interesting case omitted in Section 5, namely $k = h$ and $c = d$ in (8). This is the optimality condition for the *recommended* move c. It states that an *optimization* following move c, given the current knowledge about the other player as represented by the parameter $\sigma_h c$ of the variable $v(h, \sigma_h c)$, will not a give higher payoff to the player than when following the recommendation as expressed by $u(\sigma_h c)$ in (6). In other words, the player may choose to follow the move recommendation c now, but henceforth ignore all future recommendations and the associated Bayesian update about the other player's behavior. Intuitively, a player cannot gain by ignoring information. Formally, the claim requires that (8) is true in a bottom-up induction, starting with the information sets that are closest to the leaves. We merely prove the notationally much simpler analog for the strategic-form incentive constraints (1). The first of the inequalities (1) obviously imply for any strategy k of player 1 that

$$\sum_i \sum_j Z_{ij}\, a_{ij} \geq \sum_i \sum_j Z_{ij}\, a_{kj} = \sum_j \Big(\sum_i Z_{ij} \Big)\, a_{kj}\,. \tag{10}$$

The left hand side in (10) is the aggregate payoff to player 1 in the correlated equilibrium Z. This is the analog to the aggregate own payoff $u(\sigma_h c)$ when following the move recommendation c, and using all future information, according to (6). The right hand side in (10) is the payoff for any own strategy k when using the *marginal* probabilities $\sum_i Z_{ij}$ for the strategies j of the other player. These marginal probabilities are the analog to the entries in row $\sigma_h c$ of the correlation plan z when planning an optimized response, given the limited knowledge at the time of getting the recommendation c at h. So it is indeed optimal to follow recommendations in the future, and use the additional knowledge about the other player gained from these. □